A CHEMICAL BIOLOGY STUDY OF HUMAN EMBRYONIC STEM CELL PLURIPOTENCY AND DIFFERENTIATION

BY

YIJIE GENG

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Cell and Developmental Biology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2015

Urbana, Illinois

Doctoral Committee:

Professor Andrew S. Belmont, Chair
Professor Jie Chen, Director of Research
Professor Brian C. Freeman
Associate Professor Jing Yang

ABSTRACT

Human embryonic stem cell (hESC) research provides promising prospects for the future of regenerative medicine. To fully realize its potentials, we not only need to further our understandings toward the molecular mechanisms governing pluripotency and differentiation, but also to find new tools to manipulate the processes of directed hESC differentiations. One of the best tools for the purpose of probing and manipulating biological pathways at the molecular and cellular level is small molecules. High-throughput screening (HTS) of combinatorial chemical libraries is widely used for the search of such bioactive small molecules. My dissertation has focused on using HTS to identify novel small molecules that regulate hESC pluripotency and differentiation. Using a HTS system based on a novel design, I have screened over 170,000 small molecules (Chapter 2), and have studied three hit molecules in detail (Chapters 2, 3, and 4, respectively).

The first molecule was named Displurigen (Displg for short) for its ability to disrupt hESC pluripotency. Using a biotinylated version of Displg as a probe, I pulled down and identified HSPA8 (also known as HSC70, the constitutively expressed member of the 70 kd heat-shock protein family) as its biological target, and confirmed this finding with both loss-of-function and gain-of-function assays. I showed that HSPA8 helps maintain hESC pluripotency by interacting with OCT4 protein and facilitating its binding to DNA. This study identified HSPA8 as a novel regulatory component of the hESC pluripotency network, and was the first study to demonstrate the direct involvement of a

chaperone protein in pluripotency regulation. I describe this work in Chapter 2 of my dissertation.

The second molecule was named Mesendogen (MEG for short) for its ability to dramatically improve the efficiencies of hESC differentiation towards the mesoderm and definitive endoderm (DE) lineages. Cell-replacement therapies require differentiated cells of high purity, yet current human pluripotent stem cell differentiation protocols remain to be improved for such purposes. I attempted to enhance the commonly used growth factor-driven mesoderm and DE differentiation protocols using hit molecules acquired from the HTS, and identified MEG as a potent enhancer of growth factor-induced mesoderm and DE differentiations. Addition of MEG to mesoderm and DE differentiation cultures dramatically enhanced both their efficiencies to near homogeneity (≥ 85%). Using target identification techniques I identified transient receptor potential cation channel, subfamily M, member 6 (TRPM6) as the biological target of MEG. I then showed that MEG most likely functions by inhibiting the magnesium import activity of the TRPM6/TRPM7 channel complex, the major cellular channel of magnesium import. This study describes a robust method for the highly dependable productions of nearly homogeneous mesoderm and DE progenitors, and also reveals for the first time that TRPM6/TRPM7 channel activity and magnesium homeostasis may be involved in the regulation of early mesoderm and definitive endoderm specifications. This work is described in Chapter 3 of my dissertation.

The third molecule was named Lymphgen 1 for its activity in inducing the self-organization of a putative primitive lymphatic plexus-like structure, in the shape of a balloon, out of 2-D hESC cultures and without the addition of any growth factors. Its structural analog, Lymphgen 2, was analyzed in parallel. Recently, several studies reported self-organization events of hESCs that give rise to three dimensional (3-D) organoid structures, all of which required pre-designed programs of external guidance such as growth factors during their differentiations. Inspired by these findings, I hypothesized that by applying an appropriate stimulus such as a small molecule inhibitor at the onset of differentiation, it may be possible to trigger an intrinsic developmental program embedded in hESCs that does not require any external guidance to progress into a self-organizing structure. To find such a stimulus, I systemically examined my hit compounds for their abilities to trigger self-organizing events, and found compounds Lymphgen 1 and 2 (collectively referred to as Lymphgens hereafter) that trigger an unguided and spontaneous self-organizing event which gives rise to a balloon-like organoid structure. Gene expression analyses, functional assays, and morphological studies demonstrated that this self-organizing event may have recapitulated the *in vivo* human lymphatic morphogenesis program. Using this system, I unveiled the unguided emergence of a DiI-Ac-LDL^{+}VE-cadherin^{+}CD31^{+}CD34^{+}KDR^{+}CD43^{-} endothelial progenitor population, and identified a previously unknown VEGFR3^{+} progenitor population which may be responsible for human embryonic lymphatic development. This system provides a rare opportunity to visualize a truly spontaneous human developmental program *in vitro*. This work is described in Chapter 4 of my dissertation.

In summary, my dissertation described the development of a human pluripotent stem cell-based HTS platform that was built upon a unique and novel approach which was eventually proven to be highly effective. It also described the discoveries and detailed studies of three hit molecules found in the HTS. Through comprehensive studies of these molecules, I unveiled novel regulatory mechanisms at the molecular level that govern hESC pluripotency and differentiation. Moreover, using these small molecules as tools, I significantly improved upon current protocols for both guided lineage-specific differentiations and spontaneous organoid self-organizations of hESCs. My thesis work has been dedicated to advancing the field of hESC research through the use of chemical biology.

TABLE OF CONTENTS

CHAPTER 1: GENERAL INTRODUCTION

Background and Significance of Human Embryonic Stem Cells

Human embryonic stem cells (hESCs) are pluripotent stem cells derived from the inner cell mass of preimplantation human blastocysts (Thomson et al., 1998). Two major characteristics of hESCs are its ability to proliferate indefinitely in culture while maintaining an un-transformed (karyotypically normal) state (self-renewal), and its ability to differentiate into tissues of all three embryonic germ layers including mesoderm, endoderm, and ectoderm (pluripotency).

Originally the undifferentiated growth of hESCs was supported by the presence of mouse embryonic fibroblasts (MEFs) as feeder cells (Thomson et al., 1998). A feeder-free condition was later developed in which hESCs were cultured on Matrigel coated surfaces in medium conditioned by MEFs (Xu et al., 2001), followed by the discovery of a feeder-independent culture system including the use of a medium with a fully defined chemical composition (Ludwig et al., 2006). In all the above culture methods, undifferentiated hESCs typically grow as colonies with tightly packed cells and well-defined boarders. Differentiated cultures, on the other hand, usually display colonies and cells that lack the compact morphologies described above (Thomson et al., 1998).

From a clinical perspective, currently hESCs and human induced pluripotent stem cells (hiPSCs) and their differentiated progenies provide the best hopes for treating various diseases that require tissue replacement therapies such as type 1 diabetes, heart attack, stroke, neurodegenerative diseases, spinal cord injury, and so on. The prospect of

using hESCs and hiPSCs to generate functional organs for transplantation could also be made reality in the future.

From a basic research perspective, hESCs represent an easily accessible experimental system for the study of early human development, as availability of human embryos and fetuses are severely limited due to ethical reasons. Disease-specific hESCs and hiPSCs are now widely used as models for human genetic and developmental disorders (Wu and Hothedlinger, 2011).

Molecular mechanisms governing hESC pluripotency and differentiation

External factors and internal regulatory networks work together to maintain hESC pluripotency in the culture system. External pluripotency maintaining factors include cell culture substrates such as Matrigel and MEF feeders, and medium components such as basic nutrients and growth factors (Ludwig et al., 2006; Thomson et al., 1998). The internal pluripotency regulatory factors known to date are mostly transcription factors and epigenetic factors. Transcription factors OCT4, NANOG, and SOX2 are thought to form a core regulatory circuit which maintains hESC pluripotency (Boyer et al., 2005). Many other transcription factors form a supplementary network which connects to and supports the integrity of the core regulatory circuit (Ng and Surani, 2011). As pluripotent hESCs possess a distinct epigenetic state compared to differentiated cell types (Gifford et al., 2013), not surprisingly, epigenetic regulating factors such as chromatin remodeling complexes, DNA modification factors, and histone modification factors are also crucial to the maintenance of pluripotency (Hu and Rosenfeld, 2012). My research has unveiled

a member of the heat-shock protein family as a novel pluripotency regulatory factor, which is detailed in Chapter 2 of this dissertation.

Disruption of the pluripotency maintenance network leads to differentiation. Disruption of pluripotency by itself generally does not lead to lineage specific differentiations; rather, loss of pluripotency followed by an un-guided differentiation protocol such as the embryoid body (EB) method usually leads to the random production of multiple types of differentiated progenies at the same time. To produce specific types of differentiated cells, the differentiation processes must be precisely guided. This is generally achieved by the supplementation of external factors or direct manipulation of internal factors that regulate specific differentiation pathways. Currently the majority of regulatory pathways targeted for directed differentiation of hESCs have been discovered from developmental biology studies using animal models.

Small molecules as tools for hESC research

The term "small molecules" commonly refers to organic compounds with molecular weights lower than 900 Daltons (Macielag, 2012), as 900 Daltons is believed to be the upper molecular weight cutoff for a molecule to be able to rapidly diffuse across cell membranes. In the field of basic and clinical biomedical research, the term "small molecules" usually specifically refers to molecules with the capabilities of regulating biological processes. A more restricted use of the term refers only to molecules that binds to specific biopolymers, such as proteins or nucleic acids, and exert their biological

functions through these interactions. Most of the currently available drugs are small molecules (Macielag, 2012).

Advantages of applying small molecules for biological studies include their high potency, robustness, and ease of application. In general, nano-molar to micro-molar concentrations of a small molecule is enough to effectively and specifically inhibit its biological target(s). Small molecules can be readily applied for clinical use as drugs without the need for extensive optimizations of clinical procedures.

One disadvantage is that some small molecules may have more than one biological target, which can lead to misinterpretation of experimental results; other experimental methods must be used in parallel to confirm results acquired using small molecules. Another disadvantage is that we don't always know the biological target of a given small molecule. Overcoming this problem requires searching for their targets using target ID techniques to help us better understand the mechanism of action (MoA) of those molecules.

Small molecules have been successfully used in early studies to help maintain hESC pluripotency and to elucidate basic signaling requirements of hESC pluripotency (Gonzalez et al., 2011b; Kumagai et al., 2013; Li et al., 2007; Vallier et al., 2005; Xu et al., 2010b). An early example was the study reported by Austin Smith and colleagues, using the combination of 3 small molecules (3i) to maintain a ground state of mESC in culture that does not rely on the addition of any growth factors (Ying et al., 2008); similar

results have been reported in hESCs (Tsutsui et al., 2011). More recently, small molecules have been extensively applied to the studies of prime versus naïve pluripotency (Chan et al., 2013; Chou et al., 2008; Hanna et al., 2010; Ware et al., 2009; Zhou et al., 2010a) and induced pluripotent stem cell (iPSC) derivations (Esteban et al., 2009; Huangfu et al., 2008; Li et al., 2012; Li et al., 2009; Lin et al., 2009; Maherali and Hochedlinger, 2009; Mali et al., 2010; Moschidou et al., 2012; Wang et al., 2011; Wei et al., 2014; Yu et al., 2014; Yuan et al., 2011; Zhu et al., 2010).

Small molecules were also widely used for the study of PSC lineage-specific differentiation. BMP pathway inhibitors have been used as central components of neural ectoderm differentiation. Meanwhile a variety of small molecule inhibitors have been discovered that facilitates either general or sublineage-specific mesoderm and endoderm differentiations, such as IDE1/2 for definitive endoderm differentiation (Borowiak et al., 2009b) (although a later report challenged the effectiveness of this molecule (Tahamtani et al., 2013)), (-)-indolactam V for pancreatic lineage differentiation (Chen et al., 2009), GSK3b inhibitors for cardiomyocyte differentiation (Gonzalez et al., 2011a; Lian et al., 2012), and SB431542 for differentiation into mesenchymal progenitors (Chen et al., 2012; Mahmood et al., 2010).

To expand our current understandings of hESC pluripotency and differentiation, there is a dire need for new tools that can probe the hESC system in novel angles in the hope that previously unknown molecular mechanisms that govern pluripotency and differentiation can be revealed. It is also crucial to develop improved methods of lineage

specific differentiation for hESCs in order to achieve its full potential in regenerative medicine. Small molecules represent an ideal candidate tool for probing the molecular pathways of pluripotency and differentiation, and for improving the current differentiation methods, as will be discussed in the following sections.

Goals of my dissertation

My dissertation research has been dedicated to advancing hESC research using chemical biology. My work has been focused on: 1) using high-throughput screening (HTS) to find novel chemical compounds that regulates hESC pluripotency and differentiation; 2) finding the biological targets of those molecules using target identification (target ID) techniques to further our understandings toward the molecular mechanisms of hESC pluripotency and differentiation; and 3) deriving new and improved methods for hESC fate manipulation using the hit compounds.

Using high-throughput screening to discover novel small molecules that modulate hESC fate

The first goal of my dissertation was to use high-throughput screening (HTS) to identify novel chemical compounds that can be used to manipulate hESC fate. HTS is a drug or small molecule discovery method widely used in the field of biomedical sciences and the pharmaceutical industry. By using automated machinery and data processing, HTS allows researchers to conduct tests of a large number of chemicals in a short time frame to rapidly identify bioactive compounds. HTSs give rise to "hit" compounds. Validated hit compounds are generally referred to as "lead" compounds, meaning that

these compounds have confirmed biological activities but may require further modification and optimization of their structures to enhance their targeting efficiencies.

HTS searches bioactive small molecules from compound "libraries". A compound library may consist of compounds acquired from combinatorial chemistry (these libraries are therefore referred to as "combinatorial libraries"), natural products, collections of small molecules with known biological functions, compounds synthesized based on rational drug designs, or mixtures of the above mentioned sources.

Multiple drug candidates have been discovered through the use of HTS. Some examples of such discovery are the chemokine receptor antagonist maraviroc (Selzentry/Celsentri; Pfizer) and the thrombopoietin receptor agonist eltrombopag (Promacta/Revolade; GlaxoSmithKline) (Macarron et al., 2011).

HTSs have also been used for the discovery of novel small molecule compounds that regulate pluripotent stem cell (PSC) fate. Candidate platforms (in this case cells) that can be used for screening pluripotency regulators are the pluripotent stem cell lines, including mouse embryonic stem cells (mESCs), mouse embryonal carcinoma cells (mECCs), human embryonic stem cells (hESCs), and human embryonal carcinoma cells (hECCs). The more recent advances in the field of induced pluripotent stem cell (iPSC) technology made possible to use human iPSCs (hiPSCs) as the screening platforms as well. Disease-specific hiPSCs are especially useful for selecting drug candidates for the treatment of those diseases.

Earlier attempts of using HTS for the search of pluripotency regulators mainly focused on using mouse embryonic stem cells (mESCs) and mouse embryonal carcinoma cells (mECCs) as screening platforms. Compared to hESCs, mESCs and mECCs were easier to handle, cheaper to maintain, and more robust at staying in a pluripotent state. An example of compounds found in these studies was SC1, a small molecule that helped maintain mESC pluripotency, discovered in a mESC based screening by plating mESCs under differentiation conditions and selecting for small molecules that helped maintain pluripotency (Chen et al., 2006b). Another example was the compound TWS119, a small molecule that induced neuronal differentiation of mESCs, discovered using a HTS platform based on the use of mEC cell line P19 (Ding et al., 2003).

More recently several groups have adopted hESCs as HTS platforms. Desbordes et al. screened 2880 compounds using H9 hESCs as the screening platform and OCT4 immunostaining as readout for compounds that disrupt pluripotency (Desbordes et al., 2008). Xu et al. used HUES9 hESCs as their screening platform and ALP staining as readout for compounds that maintained cell survival and pluripotency (Xu et al., 2010b).

Despite these efforts, the problems of high cost and high levels of experimental variations that are associated with using hESCs as a HTS screening platform still remain. Researchers are now searching for ways to bypass the flaws of hESC-based HTSs. A recent study by Caron et al. attempted to solve this problem by generating a hESC line that's genetically abnormal and have characteristics that make it suitable for HTS. They

have yet to perform an actual HTS using this cell line (Caron et al., 2013). To date, there have been no reports of using human embryonal carcinoma cells (hECCs) as HTS platforms.

To overcome the limitations of using hESC for HTS, I attempted a new approach by developing a NTERA-2 hECC-based screening system for HTS. Human embryonal carcinoma cells (hECCs) are pluripotent stem cells derived from human teratocarcinomas and are considered to be the malignant counterparts of hESCs. Compared to hESCs, hECCs usually possess abnormal karyotypes and characteristics of transformed cells, meaning they are in general more robust at maintaining a pluripotent state in culture, require less care for routine maintenance, and can be maintained in simple and cheaper culture mediums without the addition of expensive growth factors. Consequently, results acquired from studies using hECCs were believed to be more stable and readily reproducible; in fact, a hECC line named 2102Ep was recommended to be included as a "reference cell line" in all PSC-related studies so that results acquired from different PSC lines maintained by different labs could be compared (Josephson et al., 2007).

I chose to use the hECC line NTERA-2 to construct my HTS platform. NTERA-2 is a clonal subline of TERA-2, one of the first established hECC lines (Andrews et al., 1984). NTERA-2 cells are able to differentiate into all three germ layers *in vivo* in the form of teratocarcinomas. In culture these cells differentiate in response to several differentiation inducers, most notably retinoic acid (Andrews, 1984), hexamethylene bis-acetamide (HMBA) (Andrews et al., 1986; Andrews et al., 1990), and bone

morphogenetic protein-7 (Andrews et al., 1994). These agents also induce differentiation of hESCs (Draper et al., 2002; Xu et al., 2002; zur Nieden et al., 2005). These properties demonstrate that NTERA-2 cells share similar molecular mechanisms underlying their pluripotent state as compared to hESCs.

I then constructed a NTERA-2 reporter cell line by stable integration of an enhanced green fluorescence protein (EGFP) reporter gene driven by a 4 kb OCT4 promoter (Gerrard et al., 2005). Using this system, I successfully screened over 171,077 small molecule compounds. Details of this screen are described in Chapter 2 of my dissertation.

Elucidating biological targets of the hit compounds acquired from HTS using target identification techniques

The second goal of my dissertation was to find the biological targets of the hit molecules using target identification (target ID) techniques, in the hopes of identifying biological pathways affected by the hit molecules, and to possibly discover novel molecular mechanisms governing hESC pluripotency and differentiation. Target ID refers to the process of identifying biological targets (most often being proteins) of bioactive small molecules using affinity reagents and pull down assays. The concept of target ID is built upon the assumption that the phenotype elicited by a small molecule arises from the cellular consequence of a direct interaction between the small molecule and its target protein(s). Many target ID strategies are available and several have been applied to my thesis study.

One of the most widely applied target ID approach is the use of a chemically modified version of a hit compound to fish out its binding target(s). Biotinylated versions of hit compounds are frequently used as probes for this type of approach. After incubation with live cell culture or cell lysates, biotinylated hit compounds, together with its binding protein(s), are immobilized by streptavidin, NeutrAvidin, or avidin beads. Eluted binding protein(s) are separated by gel electrophoresis and visualized by silver staining. Protein bands of interest are analyzed by mass spectrometry (Leslie and Hergenrother, 2008). A successful demonstration of this method is described in Chapter 2 of my dissertation.

Protein arrays have also been used to identify binding partners of a small molecule *in vitro*. Using this method, labeled hit molecules are incubated with chips coated by recombinant proteins at defined positions (dots), and scanned for signals. Dots with high signals are identified as candidate targets. An advantage of this method is that small molecules can be labeled using a wide variety of tags including radio-isotopes, fluorescent molecules, biotin, and so on. The power of this method is mainly limited by its *in vitro* binding environment and the number of recombinant proteins available on the chip. This type of protein chips is commercially available, such as the Invitrogen Proteome Array.

A commercially available screening service named "KINOMEscan" provided by DiscoveRx takes advantage of ligand competition to identify possible kinase targets of a small molecule. The basic design of this screen includes binding of a library of DNA-

tagged recombinant kinases to their ligands linked to a solid support. When a small molecule of interest binds to a certain kinase, it competes with the ligand of that kinase either by direct competition at the ligand biding site or by binding to other parts of the kinase and induce conformational change of the protein, thus releasing this kinase from the solid support. Kinases released from the ligands are collected and identified by PCR using their DNA tags as templates. A demonstration of this method is described in Chapter 3 of my dissertation.

After target ID identifies the protein target(s), several techniques can be used to further validate this interaction. These include methods that directly measure the binding of small molecules to proteins such as surface plasma resonance (SPR), and methods that measure the functional or biochemical consequences of compound-protein interactions such as a method named DARTS which detects conformational changes of a target protein (Lomenick et al., 2009; Lomenick et al., 2011), or ATPase assays and kinase assays which measure changes in the enzymatic activities of the target protein. These validation methods can only be applied when a candidate target protein is already identified, and cannot be used for exploratory purposes.

Enhancing hESC differentiation toward defined lineages using small molecules

The third goal of my dissertation was to derive new and improved methods for hESC fate manipulation using hit compounds acquired from the HTS. To realize this goal, I first searched if any of the hit molecules from the screening can guide or enhance hESC differentiation toward a defined lineage.

The major challenge of hESC-based regenerative medicine is to develop methods that can produce a large quantity of differentiated progenies that consist of a highly homogeneous (in terms of cellular identities and functions) disease-relevant cell type. Earlier attempts at hESC differentiation generally adopted a three dimensional (3-D) differentiation culture system. In this method, hESCs aggregate in suspension culture and form 3-D spheres called embryoid bodies (EBs). Embryoid bodies then undergo spontaneous differentiation and form cyst-like 3-D spheres with differentiated internal tissues consisting of cells with different identities that correspond to progenies of all three germ layers. The EB differentiation method was originally derived from mouse ES cell research, in which EBs were formed by aggregating single cells in a drop of medium hanging in the lid of a culture vessel and termed "hanging drop" method. However, in the earliest examples of hESC-based EB differentiation, instead of using the hanging drop method, researchers generated EBs from individual hESC colonies by lifting them as a whole from the cultures (Itskovitz-Eldor et al., 2000b). This was because hESCs survive poorly as single cells due to anoikis (apoptosis induced by loss of cell-cell contact). EBs obtained via this method varied greatly in size, causing experimental variations. This problem was later solved by the addition of ROCK inhibitor Y-27632 as an anoikis inhibitor to improve hESC survival. With the addition of Y-27632 it was then possible to use either the hanging drop method or specially designed micro-well culture plates to aggregate EBs from single hESCs, giving rise to well controlled, uniformly sized EBs.

However, a disadvantage of this differentiation system is that cells at the outer surface of a EB can limit the contact of differentiation inducing factors presented in the culture medium to cells on the inside of the EB. Due to this reason homogeneous differentiation within EBs can be very difficult to achieve. Thus more recent protocols often enrich target differentiation lineages by replating EBs into selective two dimensional (2-D) culture systems (Levenberg et al., 2002; Zhang et al., 2001).

A better way to produce highly homogeneous differentiated progenies may be to keep the differentiation cultures two-dimensional throughout the differentiation process. The major benefit of a 2-D culture system is that it keeps all cells exposed to the same open environment and thus provides all cells with equal access to the differentiation inducing factors in the medium, creating a differentiation condition that is easy to monitor and guide. Therefore current methods of differentiation generally adopt a 2-D culture system, within which researchers can easily apply a step-wise induction protocol and guide the differentiation process through timely addition of different growth-factor combinations during the course of differentiation (Cai et al., 2007; Chen et al., 2009; D'Amour et al., 2006; Perrier et al., 2004; Wang et al., 2007; Wong et al., 2012).

I have modified existing protocols and established consistent methods for growth factor-induced 2-D differentiations of early human mesoderm and definitive endoderm (DE) progenitors. I tested hit compounds acquired from the HTS (Chapter 2) and found MEG, a small molecule that enhanced both mesoderm and endoderm differentiation

efficiencies to near homogeneity (≥ 85%). Details of this discovery are described in Chapter 3 of my dissertation.

Deriving 3-D self-organizing organoids using small molecules

Also to achieve the third goal of my thesis study, I explored the possibility of using hit compounds acquired from the HTS (Chapter 2) to trigger an intrinsic, self-organizing differentiation process that mimics an innate human developmental program from hESCs. The term "self-organization" refers to the process of spontaneous formation of a highly ordered structure from nonprepatterned elements (Sasai, 2013).

Probably one of the biggest challenges faced by the research of early human development is the lack of an easily accessible experimental model, as the availability of human embryos or fetuses are extremely limited. The establishment of the first hESC line in 1998 by James Thomson (Thomson et al., 1998) offered a possible solution to this problem. If we can manipulate hESCs into undergoing an *in vitro* differentiation program that mimic an *in vivo* human developmental processes such as organ-formation, researchers will be able to confirm the current knowledge of human development that was largely acquired through studies using animal models, and to possibly identify novel developmental pathways that are unique to human development. Furthermore, the combined use of such systems with genetically engineered hESCs or disease-specific hiPSCs will provide models for the study of human genetic or developmental disorders, providing the hope to eventually develop treatments for these conditions. The

development of such differentiation systems remained an elusive goal until 10 years after the establishment of the first hESC line.

In 2008 researchers in Yoshiki Sasai's group reported the generation of the first self-organizing cortical tissue structures from mouse and human ESCs (Eiraku et al., 2008). Using a 3-D EB culture system developed in their lab named serum-free culture of embryoid body-like aggregates (SFEB), apico-basally polarized cortical tissues were generated. In this system, both spatial and temporal specifications of the cortical organoid could be regulated using experimental methods. The same group later developed methods for the generation of self-organizing optic-cups from mESCs (Eiraku et al., 2011) and hESCs (Nakano et al., 2012), self-forming isthmic organizertissues from mESCs (Muguruma et al., 2010), and self-forming anterior pituitary from mESCs (Suga et al., 2011), using similar 3-D EB culturing approaches.

These reports spawn a series of studies using similar or modified 3-D EB culture systems, or combinations of 2-D and 3-D culture systems, to trigger various self-organizing differentiation events using mESCs or hESCs. Examples include the self-formations of organoids resembling thyroid (Antonica et al., 2012), gut (Cao et al., 2011; Ueda et al., 2010), and pancreatic islets (Saito et al., 2011; Wang and Ye, 2009).

Another unique approach has been reported individually by two groups that relied solely on 2-D culture systems for organoid formation. Researchers in James Wells's group reported the growth of gut-like organoids out of a 2-D hPSC culture system

directed by a series of growth factor combinations (McCracken et al., 2014; Spence et al., 2011; Wells and Spence, 2014). Melissa Little's group also reported budding of embryonic kidney-like structures from 2-D hESC cultures (Takasato et al., 2014). A unique characteristic of this approach is that the 3-D structures self-organized through the process of "budding" or "sprouting" out of a 2-D surface. Similar budding or sprouting processes are present throughout development and are crucial to a variety of organogenesis events.

However, all methods published to date required the guidance of externally applied growth-factor combinations. Hoping to establish an organoid self-organization protocol dictated entirely by internal programs of the cells rather than having to rely on outside stimulus (Sasai, 2013), I hypothesized that by applying an appropriate stimuli, it may be possible to trigger an intrinsic developmental program of hESCs in culture, and thus create an organoid self-formation protocol that does not require any external guidance. In other words, by giving an initial "push" to the hESC system, we may be able to jump-start a developmental program that can, all by itself, run through its course in culture without interference. I demonstrate the induction of a self-organizing organoid-formation event by small molecules in Chapter 4 of my dissertation.

CHAPTER 2: A CHEMICAL BIOLOGY STUDY OF HUMAN PLURIPOTENT STEM CELLS UNVEILS HSPA8 AS A KEY REGULATOR OF PLURIPOTENCY

Abstract

Chemical biology methods such as high throughput screening (HTS) and affinity-based target identification can be used to probe biological systems on a biomacromolecule level, thus providing valuable insights into the molecular mechanisms of those systems. Here, by establishing an innovative human embryonal carcinoma cell-based HTS platform, I screened 171,077 small molecules for regulators of pluripotency, and identified a novel small molecule Displurigen which potently disrupts hESC pluripotency by targeting heat shock 70kDa protein 8 (HSPA8), the constitutively expressed member of the 70kDa heat-shock protein family, as elucidated using affinity-based target identification techniques and confirmed by loss-of-function and gain-of-function assays. I demonstrated that HSPA8 maintains pluripotency by binding to the master pluripotency regulator OCT4 and facilitating its DNA-binding activity. This study uncovers a novel function of HSPA8 in the maintenance of hESC pluripotency, and is the first report demonstrating the direct involvement of a chaperone protein in pluripotency regulation.

Introduction

Pluripotency regulators OCT4, NANOG, and SOX2 form a core transcription regulatory network through auto- and reciprocal-activations at the transcription level, which is believed to be responsible for the maintenance of human embryonic stem cell (hESC) pluripotency (Boyer et al., 2005). At the same time, multiple protein factors

belonging to a diversity of functional categories such as transcription factors, epigenetic factors, and signaling components work cooperatively to form an expanded pluripotency factor network which supports the core pluripotency network (Boyer et al., 2005). In contrast to the well-defined core network, our knowledge toward this expanded pluripotency network, including its components, the interactions between these components, and the mechanism-of-interaction between the expanded network and the core network, remains insufficient.

Bioactive small molecules have been applied to the field of hESC research with success. Many such studies applied small molecules as modulators of lineage-specific differentiations (Borowiak et al., 2009b; Chen et al., 2009; Chen et al., 2012; Gonzalez et al., 2011a; Lian et al., 2012; Mahmood et al., 2010); other studies exploited small molecules as chemical probes to uncover novel molecular mechanisms underlying hESC pluripotency or differentiation (Chen et al., 2006a; Xu et al., 2010a; Zhu et al., 2009). High-throughput screenings (HTS) were usually conducted for the search of such molecules. If the mechanism-of-action was unknown for a given molecule, affinity-based target identification methods can be used to identify its biological target(s). These studies have been used to identify novel protein factors and to unveil previously unknown molecular mechanisms that regulate hESC fate determination (Xu et al., 2008).

Cell-based HTSs require the use of stable and robust cellular platforms. However, hESC cultures are highly sensitive to the environment and are prone to spontaneous differentiation, thus not ideal for HTS applications (Caron et al., 2013). To solve this

problem, I chose to explore an alternative source of pluripotent stem cells, human embryonal carcinoma cells (hECCs), as a platform for HTS. hECCs are pluripotent stem cells derived from human teratocarcinomas and are considered to be the malignant counterparts of hESCs. The molecular regulatory mechanism of hECC pluripotency has been shown to mimic that of hESCs (Josephson et al., 2007). Due to their cancerous nature, hECCs are significantly less prone to spontaneous differentiation and require a much less demanding culturing condition than hESCs. Consequently, results acquired from studies using hECCs have been proven to be more stable and readily reproducible compared to hESCs (Josephson et al., 2007).

Based on the novel concept of hECC-based HTS, I established a pluripotency reporter system using the hECC line NTERA-2. Using this system, I conducted a large-scale chemical screening, and found 122 small molecules that disrupt hESC pluripotency. One of these molecules, which I named Displurigen, potently disrupts hESC pluripotency by targeting heat shock 70kDa protein 8 (HSPA8; the constitutively expressed member of the 70 kd heat-shock protein family), as discovered using affinity-based target identification methods and through functional validations. I demonstrated that HSPA8 helps maintain pluripotency by direct-binding to the OCT4 protein and facilitating OCT4 binding to DNA.

Results

Establishment of a NTERA-2 cell-based pluripotency reporter system

To avoid high experimental variations associated with hESC-based HTS platforms, I established a HTS plarform using the hECC line NTERA-2. NTERA-2 is a clonal subline of TERA-2, one of the first established hECC lines (Andrews et al., 1984). NTERA-2 cells are able to differentiate into all three germ layers *in vivo* in the form of teratocarcinomas. In culture these cells differentiate in response to several inducers of differentiation, most notably retinoic acid (RA) (Andrews, 1984), hexamethylene bis-acetamide (HMBA) (Andrews et al., 1986; Andrews et al., 1990), and bone morphogenetic protein-7 (Andrews et al., 1994); these agents also induce differentiation of hESCs (Draper et al., 2002; Xu et al., 2002; zur Nieden et al., 2005). Furthermore, basic fibroblast growth factor (bFGF) helps maintain the pluripotent state of both NTERA-2 cells and hESCs (Andrews et al., 1984; Thomson et al., 1998). The similarities in their responses toward external factors indicate that NTERA-2 cells and hESCs have very similar molecular mechanisms governing their respective pluripotent states.

An NTERA-2 cell-based reporter cell line was constructed by stable-integration of an enhanced green fluorescence protein (EGFP) reporter gene driven by a 4 kb OCT4 promoter (Gerrard et al., 2005). EGFP positive cells were selected by flow cytometry after drug selection. The selected stable cell line nearly homogeneously expressed EGFP and maintained this expression after multiple (over 20) rounds of passaging (Figure 2.1A). This reporter line is hereafter referred to as NTERA-2-OP4k, with "OP4k" referring to the 4 kb OCT4 promoter. To test the response of this reporter system to known external

regulators of pluripotency, NTERA-2-OP4k cells were cultured at low density and incubated with known inhibitors of pluripotency including RA and HMBA, and with an enhancer of pluripotency, bFGF, for 6 to 7 days. ROCK inhibitor Y-27632 was also tested as a cell survival enhancer (Watanabe et al., 2007). RA and HMBA dramatically reduced cellular EGFP signal, as shown by fluorescent imaging and FACS analysis (Figures 2.1B and 2.1C); this was accompanied by a dramatic downregulation of the endogenous protein expression of OCT4, shown by Western blotting (Figure 2.1D). On the other hand, in bFGF- and Y-27632-treated cells, the EGFP expressions and the endogenous expressions of OCT4 remained largely un-altered, as compared to the DMSO control (Figures 2.1B-2.1D). These results demonstrate that the EGFP expression of NTERA-2-OP4k reporter cells is regulated by known regulators of pluripotency in the same way as does the expression of the endogenous OCT4, and indicate that the OCT4-EGFP reporter signal faithfully reflects the endogenous expression of OCT4 and the pluripotency state of the cells.

NTERA-2-OP4k cell-based HTS identified small molecule regulators of pluripotecy

Before applying the NTERA-2-OP4k reporter system to HTS, I first verified whether the responses of the OCT4-EGFP reporter could be accurately measured under a HTS condition. For this test, NTERA2-OP4k cells were seeded at low densities (0.1 – 0.25 $\times$ 10^5 cells/cm^2) in 96-well plates. Pluripotency and survival regulators including RA, HMBA, bFGF, and Y-27632 were added immediately following cell plating. Cells were incubated for 6 to 7 days without medium change, and then scanned for EGFP signal using a fluorescent plate reader. As a result, cells treated with pluripotency

inhibitors RA and HMBA showed significantly decreased EGFP signals, while bFGF- and Y-27632-treated cells maintained high levels of EGFP signals, as compared to the DMSO control (Figure 2.2A). This result is consistent with previous results obtained by fluorescent imaging and FACS analysis under non-HTS settings (Figures 2.1B and 2.1C), thus confirming the validity of using NTERA-2-OP4k as a pluripotency reporter line for HTS.

A total of 171,077 small molecules with diverse chemical structures were screened using the NTERA-2-OP4k reporter cell-based platform. These molecules were hosted by the UIUC HTS facility in 4 chemical libraries. For the screen, NTERA-2-OP4k cells were seeded at low densities in 96- or 384-well plates and incubated with screening compounds and DMSO controls (included in all screening plates) added using pin-tools immediately after seeding, at final concentrations described in Materials and Methods. Cells were kept in a 37°C incubator for 6 to 7 days without medium change, and then scanned for EGFP signal intensities using a fluorescent plate reader. Wells with EGFP intensities lower than 70% of the average EGFP intensity of the DMSO controls (average level of background signals subtracted before comparison) in the same plates were identified as candidate hits.

I expect two types of compounds to produce false positives by reducing EGFP signal intensities in the treated wells without inducing cellular differentiation. Those include cytotoxic or apoptosis-inducing compounds that caused cell death (Figure 2.2B: "cell death"), and compounds that inhibited cell proliferation (Figure 2.2B: "reduced

growth"). A compound that fell into either of the two categories would have reduced the total number of cells in the treated well and thus able to lower its EGFP signal intensity without actually disrupting the pluripotent status of cells in that well. A real hit compound, however, would have downregulated EGFP expression on a cellular level (Figure 2.2B: "primary hits"). To distinguish between the actual hit compounds and compounds that generated false positive signals, I inspected all wells that showed significant EGFP signal reduction under a fluorescent microscope immediately after the fluorescent plate reader scan. Based on the criteria descried above, I identified the wells in which EGFP signals were reduced on a cellular level as true primary hits. As a result, while 2865 (1.7% of all compounds screened) compounds were identified as candidate hits based on their reduced EGFP signals, only 214 (0.13% of all compounds screened) of those compounds were verified as primary hits (Figure 2.2C).

I then validated the pluripotency-disruption activities of the primary hits in hESCs. Primary hit compounds were cherry-picked from the original screening plates and applied to hESCs at concentrations close to their concentrations used in the HTS (Materials and Methods). H1 and H9 hESCs were seeded as colonies and incubated with the compounds under the feeder-independent pluripotency culture condition (Materials and Methods) for 4 - 6 days; medium was changed and compounds replenished every day or every other day. Cultures were monitored every day for loss of compact colony morphologies as signs of differentiation. Compounds that failed to disrupt hESC colony integrity at the concentrations tested were examined again at 5 - 10 fold increased concentrations under the same condition. Finally, a total of 122 compounds were found to disrupt hESC colony

integrity (Figure 2.2C). These compounds are hereafter referred to as the hESC-verified hits, or verified hits.

From the verified hits, I selected the ones that showed strong potencies (effective at relatively low concentrations) at inducing hESC differentiation, acquired those compounds in-bulk from commercial sources, and further verified their activities at the molecular level. Various techniques including Western blotting, qPCR, immunostaining, and FACS analysis were used to examine changes in the expression of pluripotency markers upon treatments of those compounds. Out of the 30-plus compounds tested in detail, 29 compounds efficiently downregulated the expressions of pluripotency markers in H1 and H9 hESCs (data not shown).

Compound Displurigen potently disrupts hESC pluripotency

Among those 29 compounds, compound NSC375009 from the NCI library was one of the most potent. I hereafter refer to this compound as Displurigen (Displg for short) for its property as a "rea*gen*t that *dis*rupts *pluri*potency" (Figure 2.3A). When applied at 5 μM for 6 days, the colony integrity of hESCs was completely disrupted, as evidenced by cells within the colonies spread out and migrated away from each other (Figure 2.3B). Protein levels of OCT4 and NANOG were also downregulated after 6 days of 5 μM Displg treatment, yet SOX2 expression stayed roughly unchanged during this period (Figure 2.3C). FACS analysis under the same treatment condition also showed that, although the majority of cells lost OCT4 and NANOG expression on day 6, approximately 25% of the cells remained $OCT4^{+}NANOG^{+}$ (Figure 2.3D), indicating that

roughly 1/4 of the population remained pluripotent even after 6 days of Displg treatment at 5 μM.

To achieve a more robust differentiation, I tested Displg at higher concentrations. I found that a 10 μM concentration was sufficient for Displg to nearly completely annihilate the $OCT4^{+}NANOG^{+}$ population (from 88.29% to 0.67%; Figure 2.4A) after only 4 days of treatment. Western blotting showed similar results at the protein level (Figure 2.4B).

Both FACS analysis and Western blotting results showed that protein levels of the pluripotency markers OCT4, NANOG, and SOX2 remained largely unchanged after 24 hours of 10 μM Displg treatment (Figures 2.4A and 2.4B). In contrast, quantitative PCR (qPCR) results indicated that the mRNA levels of *OCT4* and *NANOG* decreased within 24 hours of 10 μM Displg treatment, and continued to decrease during the remaining course of differentiation (Figure 2.4C). These results showed that, upon Displg treatment, the changes of pluripotency marker expressions at the mRNA level preceded that at the protein level, indicating that the primary regulatory pathway affected by Displg may be related to the regulation of mRNA.

I then examined whether Displg induced lineage-specific differentiation of hESCs. qPCR results failed to show consistent elevation of any lineage-specific markers (Figure 2.5), indicating that Displg, although potently disrupts hESC pluripotency, is unable to induce lineage-specific differentiation. This result suggested that the biological target of

Displg most likely maintained pluripotency through direct-support of the core regulatory network rather than by blocking differentiation toward a specific lineage.

Displg disrupts hESC pluripotency by targeting HSPA8

To identify the biological target of Displg, I synthesized a biotinylated version of the compound by linking a linker and a biotin molecule at the -OH site of Displg (Figure 2.6A); this molecule is hereafter referred to as Displg-biotin. Displg-biotin disrupted hESC pluripotency, as shown by loss of colony integrity (Figure 2.6B) and downregulation of OCT4 (Figure 2.6C), although at a higher dosage (50 μM) compared to unmodified Displg. This result indicates that Displg-biotin reserves the bioactivity of its parent-compound Displg, although with a weaker potency. The decrease in potency of Displg-biotin may be due to the larger size of the molecule preventing its cellular uptake.

Using Displg-biotin as a probe, I performed an affinity-based pull-down experiment (Materials and Methods). Briefly, hESCs treated with Displg-biotin or DMSO control were lysed and then incubated with Streptavidin beads. Proteins bound to Displg-biotin were eluted and analyzed using SDS-PAGE followed by silver staining. Four distinct bands were consistently detected in the Displg-biotin pull-down samples but not in the control samples (Figure 2.7). Using mass spectrometry analysis, I identified the proteins contained in those bands as heat shock 70kDa protein 8 (HSPA8, also known as HSC70) isoform 1, telomere maintenance 2 (TELO2), SERPINE1 mRNA binding protein 1 (SERBP1), and nucleophosmin (NPM1), respectively (Figure 2.7).

One of the target proteins, HSPA8, caught my attention. HSPA8 is the constitutively expressed member of the 70 kd heat-shock protein family, as opposed to the stress-induced HSP70. Two previous publications (Fathi et al., 2009; Son et al., 2005) and my data (Figure 2.8) showed that HSPA8 is expressed in undifferentiated hESCs but downregulated upon differentiation, revealing a link between HSPA8 expression and hESC pluripotency. I thus hypothesized that HSPA8 may be the biological target of Displg.

Before testing this hypothesis, I confirmed the interaction between Displg and HSPA8 by applying the pull-down samples to Western blotting which showed that Displg indeed bound to and pulled down HSPA8 (Figure 2.9A). I also performed an *in vitro* ATPase assay using recombinant HSP70 which is structurally highly similar to HSPA8 (Stricher et al., 2013), and showed that Displg inhibited the ATPase activity of HSP70 with an IC_{50} of 225 μM (Figure 2.9B). These results confirmed HSPA8 as a target of Displg.

I then conducted loss-of-function experiments to examine whether HSPA8, and not the other proteins bound by Displg, was the true biological target of Displg in regard to its activity in disrupting pluripotency. Knock-down of HSPA8 in hESCs using shRNAs disrupted pluripotency, as shown by loss of colony integrity (Figure 2.10A) and downregulation of pluripotency markers OCT4, NANOG, and SOX2 (Figure 2.10B). Other commercially available HSP70/HSPA8 inhibitors including Apoptozole (Cho et al., 2011) (20 μM) and VER155008 (Massey et al., 2010) (20 μM) also disrupted hESC

pluripotency, although less potently compared to Displg (Figure 2.10C). On the other hand, knock-down of the other protein targets pulled down by Displg, including NPM1, SERBP1, and TELO2, failed to disrupt hESC pluripotency (Figures 2.11). These results demonstrated that HSPA8 is most likely to be the biological target of Displg which is responsible for maintaining hESC pluripotency.

To further verify HSPA8 as the primary target of Displg in pluripotency maintenance, I attempted to rescue the pluripotency-disruption effect of Displg by overexpressing HSPA8 in hESCs. I first overexpressed wild-type HSPA8 protein, which rescued the effect of Displg, as shown by the elevated OCT4 expression (Figure 2.12A). The efficiency of this rescue was modest, which may be attributed to the fact that overexpression of wild-type chaperone proteins has been known to be harmful to cells (Freeman et al., 1995). To achieve a more robust rescue, I overexpressed a chaperone-dead mutant form of HSP70 protein named HSP70-AAAA (Freeman et al., 1995), in which the last 4 amino acids were mutated from EEVD to AAAA. Indeed, overexpression of HSP70-AAAA gave rise to a more robust rescue of Displg as demonstrated by the pluripotency marker expressions (Figure 2.12B). These results confirmed HSPA8 as the biological target of Displg, and further verified the function of HSPA8 in pluripotency maintenance.

HSPA8 maintains pluripotency by facilitating the DNA-binding activity of OCT4

Heat shock proteins, as molecular chaperones, are involved in a variety of cellular processes such as maintenance of protein stability (Wegele et al., 2004) and mediating

the assembly and disassembly of protein complexes (Ellis, 2007). More recently researchers discovered that heat shock proteins also play an important role in regulating the dynamics of protein-DNA association events such as transcription (Freeman and Yamamoto, 2002; Muller et al., 2004; Shaknovich et al., 1992; Stavreva et al., 2004; Walerych et al., 2004). I thus attempted to examine the effect of HSPA8 on the activity of the core transcription factor network of pluripotency. For simplicity, I focused my study on the interaction between HSPA8 and OCT4 which is the master regulator of pluripotency. Indeed, co-IP experiments showed that HSPA8 bound to OCT4, and that this interaction was disrupted by Displg (Figure 2.13).

My data showed that the downregulation of core pluripotency factors at the mRNA level preceded the downregulation of those factors at the protein level (Figure 2.4); specifically, mRNA levels of the pluripotency factors were downregulated within the first 24 hours of Displg treatment (Figure 2.4C), whereas their protein levels remained largely unchanged after 24 hours of treatment (Figures 2.4A and 2.4B). This distinct dynamic indicated that HSPA8 most likely regulates the core pluripotency network at the mRNA level.

I thus hypothesized that HSPA8 maintains pluripotency through facilitating the transcription-activation activity of OCT4. To test this hypothesis, I examined whether HSPA8 was required for OCT4 protein to efficiently bind to DNA. I first tested this potential regulatory mechanism using electrophoretic mobility shift assay (EMSA), and discovered that Displg decreased the binding affinity of OCT4 to its DNA-binding motifs

found in the *OCT4*, *NANOG*, and *SOX2* promoters (Catena et al., 2004; Chew et al., 2005; Rodda et al., 2005). Two distinct bands (A & B; Figure 2.14A) were detected in EMSA; a supershift assay using anti-OCT4 antibody was conducted which showed band-A as the band representing OCT4-binding to the DNA probes (Figure 2.14B). This data indicate that the activity of HSPA8 was necessary for OCT4 protein to efficiently bind to DNA. I then examined whether this regulatory effect depends on the presence of other protein factors by conducting an *in vitro* binding assay using recombinant OCT4 and HSP70 proteins. OCT4 alone only exhibited a weak binding affinity to DNA, whereas the addition of 1 - 10 μM recombinant HSP70 dramatically enhanced the OCT4-DNA interaction (Figure 2.15). This result suggests that HSPA8 facilitates the association of OCT4-DNA complexes in a direct fashion independent of other protein factors.

Discussion

In this study, I established a NTERA-2 hECC-based pluripotency reporter system and screened 171,077 compounds. I identified 29 bioactive small molecules that potently disrupted hESC pluripotency as evidence by changes in pluripotency marker expressions, and studied one particular molecule, Displurigen, in detail. Using target identification techniques and functional verifications I discovered HSPA8 as the biological target of Displurigen.

HTS has been used for the discovery of novel small molecule compounds that regulate pluripotent stem cell (PSC) fate; such studies require the use of pluripotent cell lines as screening platforms. Commonly available pluripotent platforms include mouse

and human embryonic stem cells (ESCs), mouse and human embryonal carcinoma cells (ECCs), and the more recently established mouse and human induced-pluripotent stem cells (iPSCs). The earliest attempts of pluripotency-regulator screenings adopted mouse embryonic stem cells (mESCs) and mouse embryonal carcinoma cells (mECCs) as their screening platforms (Chen et al., 2006a; Ding et al., 2003). More recently, several hESC-based HTSs have been reported (Barbaric et al., 2010; Desbordes et al., 2008; Xu et al., 2010a). Compared to hESCs, mESCs and mECCs are better platforms for HTS: they are easier to passage and maintain, and are less prone to spontaneous differentiation. However, to achieve deeper understandings to human stem cell pluripotency and differentiation as demanded by practical applications such as regenerative medicine requires more robust platforms for pluripotency-related HTSs. A recent study by Caron et al. attempted to solve this problem by adopting a genetically abnormal hESC line with characteristics that make it favorable to HTS; they have yet to perform an actual HTS using this model (Caron et al., 2013). Barbarics et al. chose to use unmodified NTERA-2 cells as their screening platform and conducted a morphology based screen to search for pluripotency regulators, but were only able to screen 80 compounds using their method. To date, there have been no reports of truly "high-throughput" screenings conducted using hECC-based platforms. Therefore my system represents the first hECC-based HTS platform which has been proven effective through an actual screen.

Heat-shock proteins, as molecular chaperones, are involved in a variety of cellular functions. Apart from their commonly known roles in maintaining protein stability and mediating the assembly and disassembly of protein complexes, recent studies showed that

certain chaperones, including HSPA8, also regulate the dynamic formation of protein-DNA complexes involved in biological processes such as transcription, telomere maintenance, DNA repair and DNA replication (Ahn et al., 2005; DeZwaan and Freeman, 2010). My study unveiled a novel function of HSPA8 in maintaining hESC pluripotency by facilitating the DNA-binding, and subsequently, transcription-activation activities of the master pluripotency factor OCT4. This is the first report demonstrating a direct involvement of a chaperone protein in the regulation of pluripotency.

My study also adds to the current reservoir of bioactive small molecules a novel HSPA8/HSP70 inhibitor, Displurigen. Modulation of heat-shock protein activities has been extensively studied for potential clinical applications such as cancer therapy. In contrast to normal cells, cancer cells strongly overexpress HSP70 in order to provide resistance to stresses generated by the environment, from tumorigenesis events, and during cancer therapy. This addiction to HSP70 provides the theoretical basis for its targeting in anti-cancer treatment (Goloudina et al., 2012; Jego et al., 2013). In the recent years, several novel HSPA8/HSP70 inhibitors have been discovered and studied for anti-cancer therapies (Goloudina et al., 2012). The discovery of Displurigen expands this growing list of HSPA8/HSP70 inhibitors and provides a novel drug candidate for cancer therapy.

In summary, by establishing and applying a hECC-based HTS system, my study identified a novel chemical inhibitor of HSPA8/HSP70, and unveiled a previously

uncharacterized, chaperone-related molecular mechanism for the maintenance of hESC pluripotency.

Materials and Methods

Cell culture

H9 and H1 hESC lines (WiCell Research Institute, Madison, WI) were maintained under a feeder condition or a feeder-independent condition. For the feeder condition (Thomson et al., 1998), primary mouse embryonic fibroblasts (MEFs) prepared from embryos of pregnant CF-1 mice (day 13.5 of gestation; Charles River) were cultured in Dulbecco's Modified Eagle Medium (DMEM) containing 10% FBS (Hyclone), 1% non-essential amino acids (NEAA; Invitrogen), and penicillin/streptomycin, and mitotically inactivated by gamma irradiation. H9 and H1 hESCs were cultured on irradiated MEFs in media containing DMEM/F12, 20% knockout serum replacement (KSR; Invitrogen), 4 ng/ml basic fibroblast growth factor (bFGF) (Invitrogen), 1% NEAA, 1 mM glutamine, and 0.1 mM β-mercaptoethanol. For the feeder-independent condition, hESCs were cultured on Matrigel (BD Biosciences)-coated plates in mTeSR1 medium (StemCell Technologies) as described (Ludwig et al., 2006). The experiments described in this study were conducted with H9 and H1 hESCs between passages 30 and 60. NTERA-2 cells (NTERA-2 cl.D1) were purchased from American Type Culture Collection (ATCC) and were cultured in Dulbecco's Modified Eagle's Medium (DMEM) containing 10% fetal bovine serum.

Generation and validation of NTera-2-OP4k reporter cells

The plasmid containing EGFP driven by a fragment of OCT4 promoter (~4 kb) was kindly provided by Dr. Wei Cui (Imperial College London) and was as described (Gerrard et al., 2005). We established NTERA-2-OP4k cells containing the OCT4-EGFP construct by transfecting cells using Amaxa Nucleofection System (Nucleofector Kit L; Program X-001), selecting transfected cells with G418 (500 μg/ml, 2 weeks), and enriching EGFP-positive cells using fluorescence-activated cell sorting (FACS) (Cytomation Plus, Dako).

For validation of reporter activity, NTERA-2-OP4k cells were plated at a density of $0.1 – 0.25 \times 10^5$ cells/cm^2 and incubated with RA (Sigma; 10 μM), HMBA (Sigma; 3 mM), bFGF (Invitrogen; 4 ng/ml), and Y-27632 (Calbiochem; 10 μM). After 6 - 7 days of incubation, changes in the level of EGFP signals were examined using fluorescent microscopy (Zeiss), flow cytometry (BD Biosciences LSR II), and fluorescent plate reader (Analyst HT, Molecular Devices).

Large-scale chemical screening

Large-scale chemical screening was conducted at the High-Throughput Screening Facility (HTSF) at the University of Illinois at Urbana−Champaign (http://www.scs.illinois.edu/htsf/index.html). The HTSF hosts 171,077 compounds from several compound libraries, which include the Marvel Library, the HTSF House Library, the ChemBridge MicroFormat Library, and the National Cancer Institute (NCI) library.

For large-scale chemical screening, NTERA-2-OP4k cells were trypsinized and seeded onto 96/384-well plates (at 0.1 – 0.25 × 10^5 cells/cm^2) using a WellMate Microplate Dispenser (Matrix). Compounds were added immediately after plating using a 96-well or a 384-well pin-tool. The first two and last two columns of 384-well plates and the first and last column of 96-well plates were used for DMSO treatment as negative controls. Cells were incubated for 6 - 7 days without medium change. EGFP expressions of individual wells were recorded by the fluorescence plate reader (Analyst HT, Molecular Devices). EGFP signal detected with compound treatments were compared to the DMSO control, and those with significant reduction (>30%; average level of background signals subtracted before comparison) were marked as potential hits and were subsequently inspected visually using a fluorescence microscope (Zeiss). The final compound concentrations applied to the screening plates were 5 - 10 μM (Marvel library), 500 nM - 1 μM (NCI library), 5 - 10 μM (House library), and 10 - 20 μg/ml (ChemBridge MicroFormat library).

Western blotting

Cultured cells were lysed directly by 2× Laemmli buffer (Bio-Rad), boiled for 5 min, and analyzed using SDS-PAGE electrophoresis followed by wet-transfer onto nitrocellulose membranes using a system manufactured by Bio-Rad. The membranes were blocked using blocking solution (5% BSA in Tris-buffered saline containing 0.1% Tween-20 [TBST]), and then incubated with primary antibodies, diluted in TBST, at 4°C overnight. The membranes were then washed by TBST for 3 × 5 min, and incubated with horseradish peroxidase (HRP) conjugated secondary antibodies at room temperature for

1 hr. Finally, the membranes were washed 5 min each time for 3 – 5 times by TBST and developed using Super-Signal West Pico Chemiluminescent Substrate (Pierce).

RNA extraction, reverse transcription, and quantitative-PCR

Total RNA were isolated using RNeasy mini kit (QIAGEN). cDNAs were synthesized from the purified RNAs using Reverse Transcription System (Promega). Quantitative-PCR was performed using QuantiTech SYBR Green PCR kit (QIAGEN). Signals were analyzed using the comparative C_T method, and ACTB gene was used as an internal control.

Flow cytometry

Single cell suspensions were acquired through Accutase (Invitrogen) treatment. Cells were fixed and stained using the Transcription Factor Buffer Set (BD Biosciences) following the manufacturer's instructions. Conjugated antibodies including OCT4A-Alexa647 and NANOG-PE were used. Cells were resuspended in PBS supplemented with 1% BSA and analyzed using a BD Biosciences LSR II flow cytometry analyzer and BD FACSDiva software.

Affinity chromatography, electrophoresis, and silver staining

hESCs were cultured under the feeder-independent condition with or without the presence of displurigen-biotin (10 μM) overnight, washed in PBS three times and lysed in Ice-cold RIPA buffer supplemented with protease and phosphatase inhibitors cocktail (Sigma). The lysates were then diluted two-fold by Tris buffer (50 mM Tris-HCl, pH 7.5).

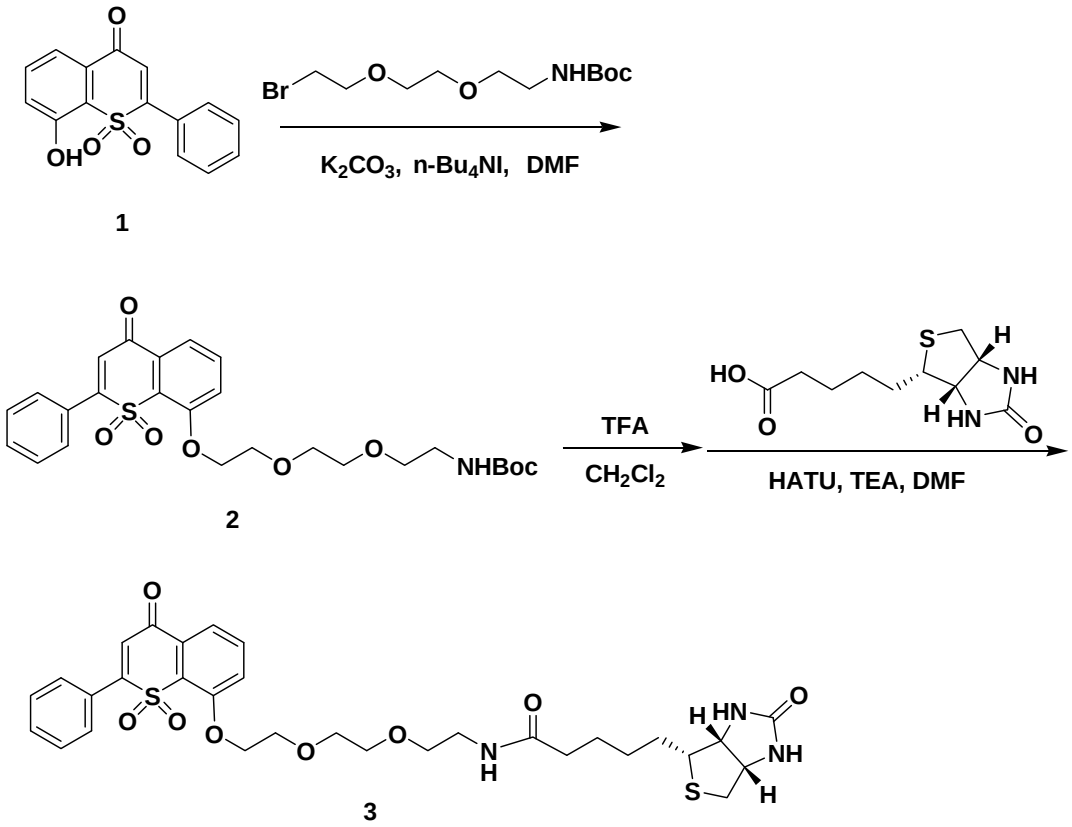
For affinity chromatography ("pull-down"), diluted lysates were incubated with streptavidin−agarose beads (Sigma) for 5 hr at 4°C. The beads were collected using centrifugation and washed with wash buffer (75 mM NaCl, 0.5 mM EDTA, 0.5% Triton X-100, 0.5% sodium deoxycholate, 0.05% SDS, 50 mM Tris-HCl, pH 7.6). The washed beads were suspended in 2× Laemmli buffer (Bio-Rad) and heated at 100°C for 5 min.

For electrophoresis and silver staining, 10 μl samples were loaded on polyacrylamide gel. Silver staining was performed using ProteoSilver Plus Silver Stain Kit (Sigma). Protein bands detected by silver staining were selectively excised for mass spectrometry analysis.

Lentivirus production and infection

For viral packaging, expression vectors were co-transfected with pCMV-dR8.91 and pCMV-VSV-G into 293T cells by $CaPO_4$ precipitation. After overnight incubation, culture medium was replaced by virus-packaging medium containing DMEM, 30% FBS, and 1 mM sodium pyruvate. Supernatants were collected 48 hr later and concentrated approximately 100× by ultracentrifuge (20,000 rpm, 1 hr). H1 and H9 hESCs were infected by virus concentrates in the presence of Polybrene (6 μg/ml), and then subjected to puromycin (0.5 μg/ml) selection for 3-5 days prior to analysis.

Synthesis of Displurigen-biotin



Synthesis of Compound 2: Compound 1 (0.0055 g, 0.019 mmol), 2-[2-(2-tertbutoxycalbonyl aminoethoxy) ethoxy]ethyl bromide (0.0071 g, 0.023 mmol), K_2CO_3 (0.0030 g, 0.022 mmol) and *n*-Bu_4NI (0.0014 g, 0.0038 mmol) were suspended in DMF (1 mL). The suspension was heated at 60 °C for 12 h. After evaporation of the solvent, crude product was purified by flash chromatography on silica gel (EtOAc : Hexanes = 1 : 1) and then EtoAc to give **2** as solid (0.0086 g, 0.0166 mmol, 86.5 %). ^{1}H NMR (300MHz, $CDCl_3$, 25 ^{o}C) δ 1.43 (s, 9H), 3.29 (m, 2H), 3.53 (t, *J* = 5.1 Hz, 2H), 3.58 (m, 2H), 3.76 (m, 2H), 3.99 (t, *J* = 4.5 Hz, 2H), 4.40 (t, *J* = 4.8 Hz, 2H), 4.98 (bs, 1H), 6.73 (s, 1H), 7.42-7.62 (m, 4H), 7.68 (t, *J* = 6.9 Hz, 1H), 7.83-7.90 (m, 3H); ^{13}C NMR (75 MHz,

$CDCl_3$, 25 ^oC) 28.4, 40.4, 69.3, 70.2, 70.4, 70.5, 71.4, 119.9, 120.4, 127.8, 129.1, 130.0, 131.6, 133.8, 154.5, 155.9, 156.6, 178.4

Synthesis of Compound 3: Deprotection of **2** (0.0086 g, 0.0166 mmol) was performed with 0.5 mL of CH_2Cl_2 and 0.2 mL TFA at 0 °C. After stirring at 0 °C for 3 h, the reaction mixture was concentrated. Toluene (0.5 mL) was added to the residue and then evaporated to remove TFA. The procedure was repeated three times. The resulting TFA salt of the deprotected amine was dissolved in DMF (1 mL) and D-biotin (0.0071 g, 0.029 mmol). HATU (0.012 g, 0.0316 mmol) was added followed by a triethylamine (25 μL, 0.18 mmol). After stirring at rt overnight, the mixture was concentrated under reduced pressure and the residue was purified by flash chromatography on silica gel ($CHCl_3$: EtOH = 10 : 1), which gave **3** (0.0077 g, 0.012 mmol, 72.1 %) as solid. ^1H NMR (300 MHz, $CDCl_3$, 25 °C) δ 1.37 (m, 2H), 1.68 (m, 4H), 2.15 (t, *J* = 7.5 Hz, 2H), 2.70 (d, *J* = 12.9 Hz, 1H), 2.88 (dd, *J* = 5.4 Hz, *J* = 12.9 Hz, 1H), 3.09 (m, 1H), 3.40 (m, 2H), 3.56 (t, *J* = 5.1 Hz, 2H), 3.67 (m, 2H), 3.78 (m, 2H), 3.98 (t, *J* = 4.8 Hz, 2H), 4.23 (m, 1H), 4.40 (m, 2H), 4.47 (m, 1H), 4.93 (s, 1H), 5.77 (s, 1H), 6.50 (t, 1H), 6.75 (s, 1H), 7.42-7.60 (m, 4H), 7.66 (t, *J* = 7.5 Hz, 1H), 7.84 (m, 3H); ^{13}C NMR (75 MHz, $CDCl_3$, 25 ^oC) 25.8, 28.3, 36.1, 39.4, 40.8, 55.7, 60.4, 62.0, 69.4, 70.2, 70.6, 70.7, 71.2, 120.0, 120.7, 128.2, 129.4, 130.3, 131.9, 134.2, 154.6, 156.7, 164.0, 173.5, 178.6; MS (ESI^+) (*m*/*z*): $[M + H]^+$ Calcd. For $C_{31}H_{38}N_3O_8S_2$: 644.2095; Found: 644.2104.

Mass spectrometry analysis

In UIUC MS facility: Gel slices were dehydrated and destained in 50% Acetonitrile + 25 mM NH4HCO3 and gently crushed using a plastic pestle inside a 1.5 ml eppendorf tube. The crushed gel was dried briefly using Speedvac (Thermo) before digestion with Trypsin (mass spectrometry grade, G-Biosciences at 1:50 w/w) in 25 mM ammonium bicarbonate. Digestion was performed using a CEM Discover Microwave Digestor (Mathews, NC) for 15 min at 75 watts, 55°C. The digested peptides were extracted using 50% acetonitrile and 5% formic acid, lyophilized and reconstituted in 5% acetonitrile and 0.1% formic, 10 µl of which was used for LC/MS analysis.

Mass spectrometry was performed using a Waters Q-ToF connected to a Waters nanoAcquity UPLC (Milford, MA). Column was a Waters Atlantis dC18 nanoAcquity UPLC column 75 µm x 150 mm (3 µm particle size) running at 250 nl/min. Gradient was from 100% A to 60% B (A: water and 0.1% formic acid; B: acetonitrile and 0.1% formic acid). Data collection was performed using MassLynx 4.1 (Waters). The top 4 intensive precursor ions from each survey scan were subjected to MS/MS by Collision Induced Dissociation (CID). The raw mass spectrometric data were processed using ProteinLynx Global Server 2.2.5 (Waters). The refined peaklists were analyzed using Mascot 2.2 (Matrix Science, London, UK) with a tolerance of ± 0.4 Da for both the precursor ions and fragment ions. Searches were carried out using the NCBI non-redundant protein database.

In UC-Riverside MS facility: Each gel band was destained (1:1 mixure of 30 mM K3Fe (CN)6 and 100 mM Na2S2O3), then dehydrated with acetonitrile, and digested with sequencing grade trypsin (10μg/mL) overnight at 37 oC. The resulting peptides were extracted and prepared as previously described (Carter et al., 2004). A nano-UPLC/MS/MS analysis was performed using combination of the nano-Acquity UPLC and Q-TOF Premier (Waters) following the method in a previous study (Rosado et al., 2010).

Embryoid body (EB) formation

H9 and H1 hESC colonies were dissociated from the culture surface by 20 min treatment of Dispase (Invitrogen). Suspended colonies were pooled by brief centrifugation (1000 rpm, 1 min), resuspended in medium containing Advanced RPMI 1640 (Invitrogen), 2% B-27 supplement (Invitrogen), and 1% Glutamax (Invitrogen), and then plated into ultra-low attachment plates (Corning). Medium was changed every other day.

In vitro ATPase assay

The HSP70 ATPase rates were determined as described (Freeman et al., 1995). In brief, ATP hydrolysis was determined by measuring the release of [^{32}P]Pi from [γ-^{32}P]ATP. The ATPase rates were calculated utilizing an average [γ-^{32}P]ATP hydrolysis rate at each time point (5, 10, 15 and 20 min) from three separate experiments after the background hydrolysis was subtracted. The data were visualized and quantified by PhosphorImager analysis (Molecular Dynamics). The effect of displurigen on the ATPase

rate of HSP70 was measured by incubating HSP70 (5 min) with varying levels of displurigen prior to the introduction of ATP.

HSPA8 overexpression plasmid construction

HSPA8-isoform-1 cDNA clone was purchased from OriGene (SC322471), and subcloned into a pSin-EF2 plasmid (Addgene; modified with a short adaptor sequence) at the MluI/NdeI site.

Immunoprecipitation

Cells treated with DMSO or Displg (100 μM, 2 hours) were lysed in ice-cold RIPA buffer (Pierce) supplemented with a protease inhibitor cocktail (Sigma). After clarification, cell lysates were pre-cleared with Protein A/G magnetic beads (Pierce) and thereafter incubated with 2 μg of anti-OCT4 antibody (Santa Cruz) and with DMSO or 1 mM Displg, respectively. After 2 hours of incubation at 4°C, Protein A/G magnetic beads were added to the cell lysates. The resulting immunoprecipitates were washed with ice-cold PBS and analyzed by Western blotting.

Electrophoretic mobility shift assay (EMSA)

EMSA analysis was performed using Buffer C nuclear extracts (10 μg) from hESCs. The extracts were incubated with poly dI-dC (Sigma) and ^{32}P-end labeled oligonucleotides: *OCT4*, *NANOG* and *SOX2* binding sequence probes along with the complimentary primers. The protein-DNA complexes were resolved by native

polyacrylamide gel (4%) electrophoresis and the dried gels were visualized using a PhosphorImager (Molecular Dynamics).

In vitro OCT4-DNA binding assay

OCT4 cDNA was cloned from the pSin-EF2-OCT4-Pur vector (Addgene) and inserted into the PET-24a protein expression vector, which contains a C-terminal His Tag. OCT4 and HSP70 proteins were expressed in Rosetta cells and captured with the Talon metal affinity resin (Clontech). The proteins were further purified using Resource Q and Superdex 200 columns (GE Healthcare). OCT4 protein binding to the *OCT4* response element was assessed by the use of EMSA analysis in the presence or absence of purified HSP70 protein.

Statistical analysis

Statistical analyses were performed using Microsoft Excel. Student's t-test was used to compare two experimental groups, assuming unequal variances. Differences are considered significant when $p < 0.05$.

Figures

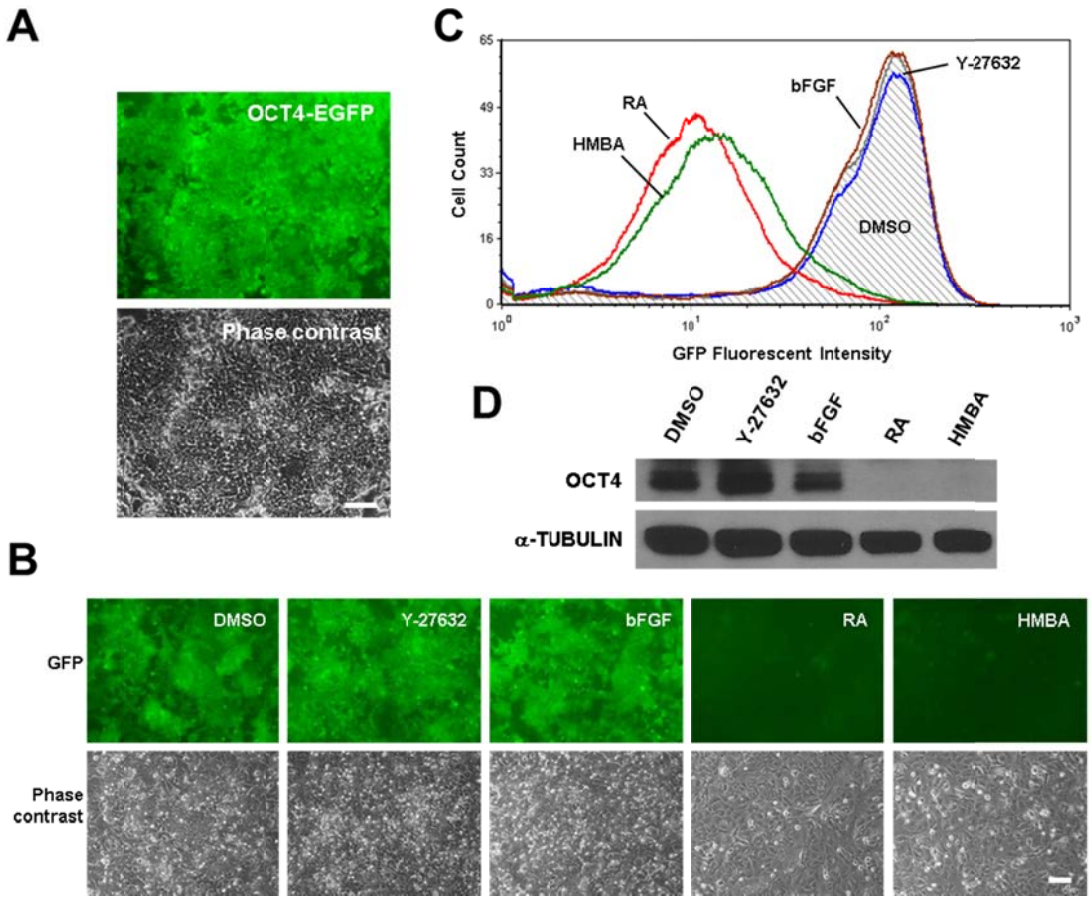
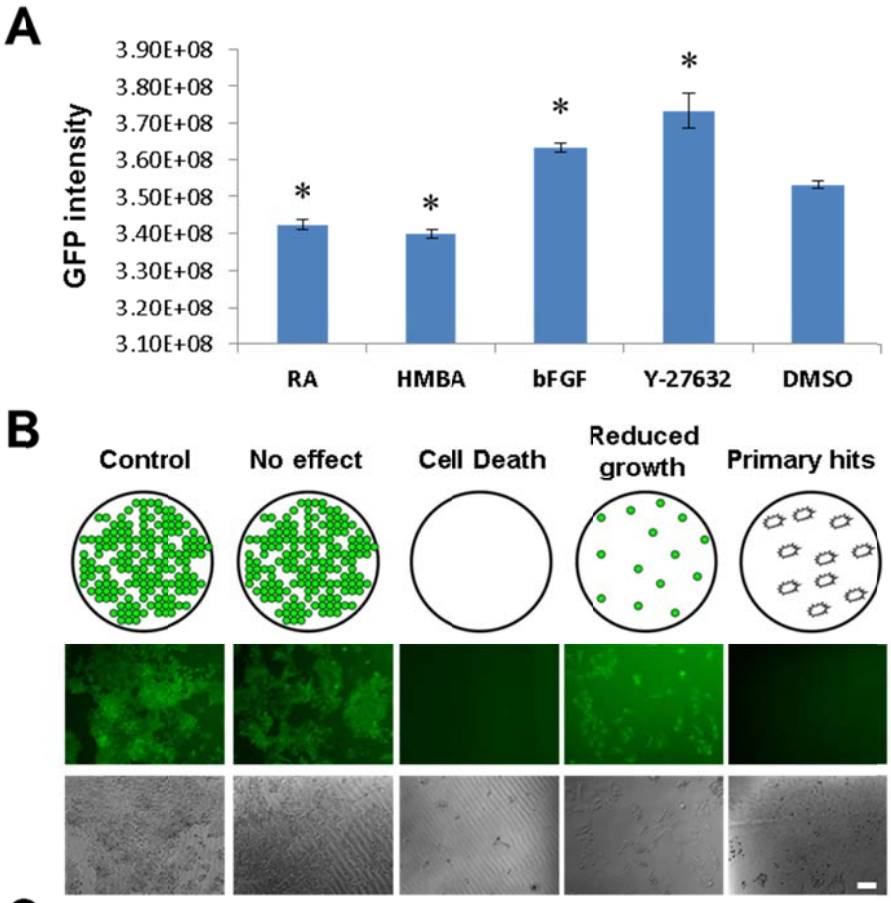


Figure 2.1. Generation and validation of NTERA-2-OP4k reporter cell line. (**A**) Fluorescent (top) and phase contrast (bottom) images of NTERA-2-OP4k cells. Scale bar: 100 μm. (**B**) Fluorescent (top panel) and phase contrast (bottom panel) images of NTERA-2-OP4k cells treated with DMSO, Y-27632 (10 μM), bFGF (4 ng/ml), retinoic acid (RA; 10 μM) and HMBA (3 mM) for 6 days. Scale bar: 100 μm. (**C**) Histogram of GFP expression in NTera-2-OP4k cells analyzed by flow cytometry. Cells were treated with DMSO, Y-27632, bFGF, RA, and HMBA for 6 days. (**D**) Western blotting of OCT4 in NTERA-2-OP4k cells treated with DMSO, Y-27632, bFGF, RA, and HMBA for 7 days. α-TUBULIN was used as a loading control.



C

Compound libraries	No. of compounds screened	Candidate hits identified by plate reader	Primary hits verified by microscopy	hESC-verified hits
Marvel library	10,300	155 (1.5%)	45 (0.44%)	25 (0.24%)
NCI library	10,277	310 (3%)	28 (0.27%)	17 (0.17%)
House library	500	300 (60%)	18 (3.65%)	5 (1%)
Chambridge library	150,000	2,100 (1.4%)	123 (0.08%)	75 (0.05%)
Total	171,077	2,865 (1.7%)	214 (0.13%)	122 (0.07%)

Figure 2.2. The strategy and general outcome of NTERA-2-OP4k cell-based HTS. (**A**) Quantification of the GFP intensity after treatment of NTERA-2-OP4k cells with DMSO, Y-27632, bFGF, RA, and HMBA for 6 days. Cells were cultured in the 96-well-plate format. The GFP intensity was determined using an Analyst HT Fluorescence Plate Reader. Each bar represents mean ± SD (error bars). n = 4 independent experiments. *: $p<0.05$, as determined by Student's t test.

Figure 2.2. (Cont.)
(**B**) Schematic representation (top panel) of several cellular outcomes of NTERA-2-OP4k cells after the chemical screening. Representative fluorescent (middle panel) and phase-contrast (bottom panel) images were taken from the actual screening and are shown for each outcome. Scale bar: 100 μm. (**C**) Number of total compounds screened and hit compounds identified after each step of screening. The hit rates (in parenthesis) were calculated based on the ratio of hit compounds and total compounds. Compounds from 4 different libraries (Marvel, NCI, House, and ChemBridge library) were screened. Due to the large number of compounds hosted by the libraries (171,077 compounds total), only one round of screening (n = 1) was performed for the entire compound collection.

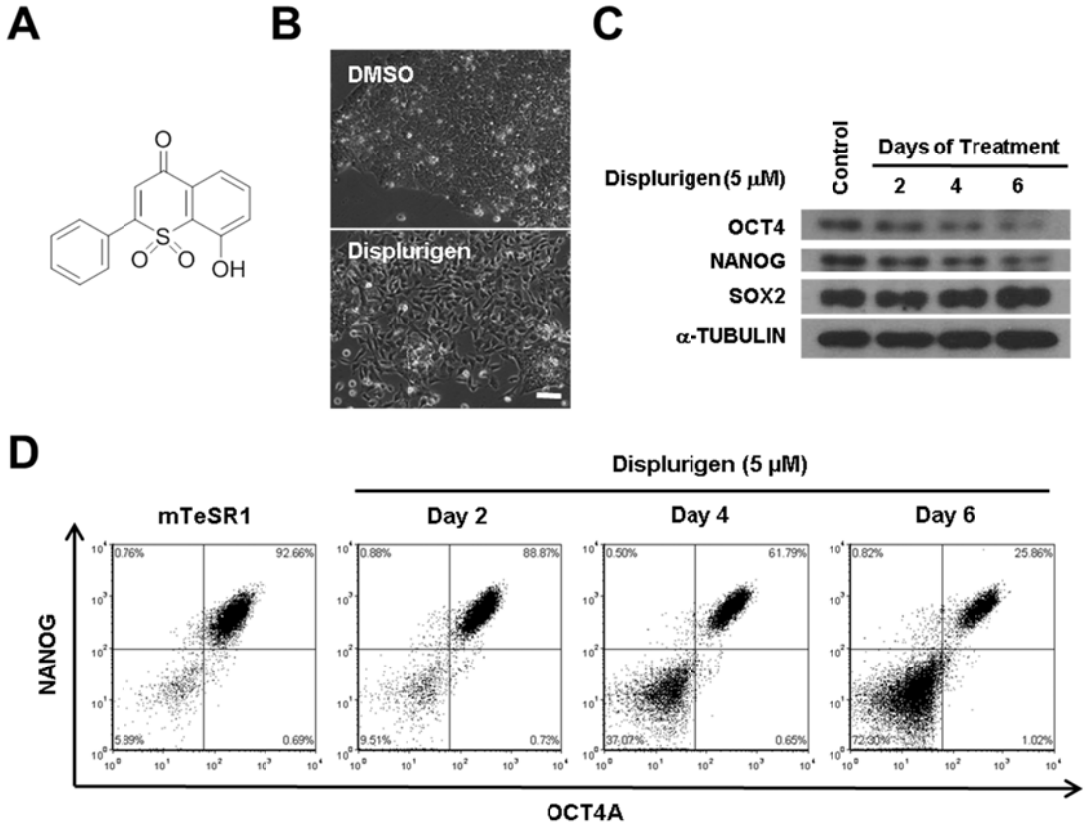


Figure 2.3. Compound NSC375009 (Displurigen) partially disrupts hESC pluripotency at 5 μM. (**A**) Chemical structure of compound NSC375009 (Displurigen). (**B**) Phase-contrast images of H9 hESCs treated with DMSO and compound NSC375009 (Displurigen) (5 M) for 2 days. Scale bar: 100 μm. (**C**) Western blotting of OCT4, NANOG, and SOX2 in H1 hESCs untreated (mTeSR1) or treated with NSC375009 (Displurigen, 5 μM) for 2, 4, and 6 days. α-TUBULIN was used as a loading control. (**D**) FACS analysis of OCT4A and NANOG in H1 hESCs untreated (mTeSR1) or treated with 5 μM NSC375009 (Displurigen) for 2, 4, and 6 days.

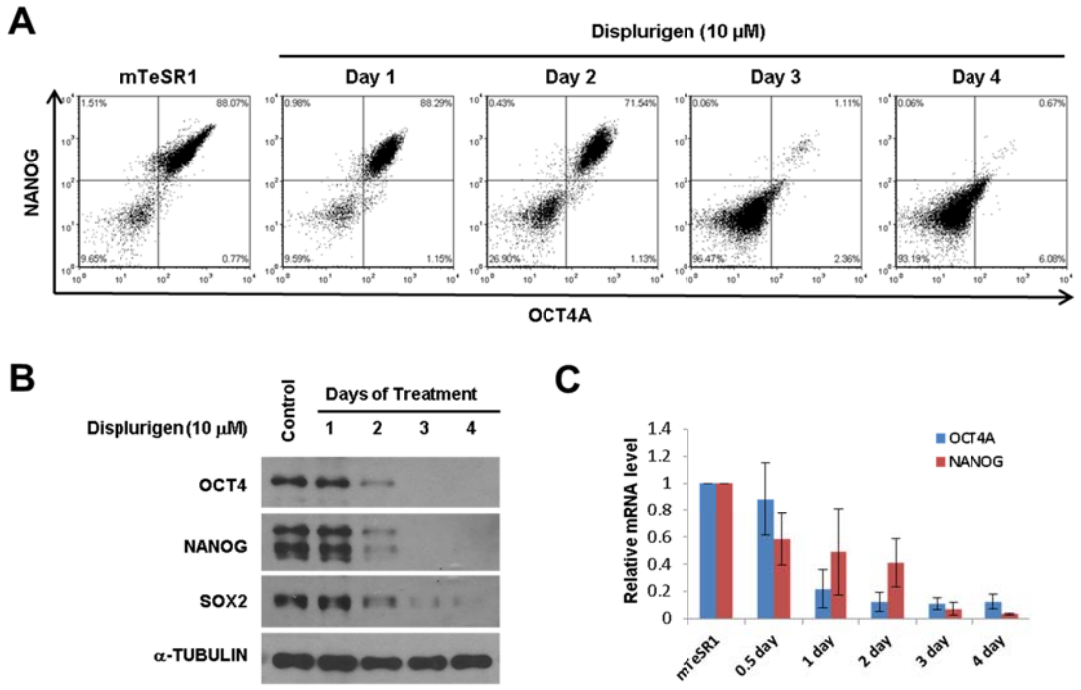


Figure 2.4. Compound NSC375009 (Displurigen) quickly and completely disrupts hESC pluripotency at 10 μM. (**A**) FACS analysis of OCT4A and NANOG in H1 hESCs untreated (mTeSR1) or treated with 10 μM NSC375009 (Displurigen) for 1 - 4 days. (**B**) Western blotting of OCT4, SOX2, and NANOG in H1 hESCs untreated (mTeSR1) or treated with 10 μM NSC375009 (Displurigen) for 1 - 4 days. α-TUBULIN was used as a loading control. (**C**) Quantitative-PCR analysis of *OCT4* and *NANOG* in H1 hESCs untreated (mTeSR1) or treated with 10 μM NSC375009 (Displurigen) for 0.5 to 4 days. All values were normalized to the level (=1) of mRNA in cells without compound treatment. Each bar represents mean ± SD (error bars). *ACTB* (*β-actin*) was used as a loading control.

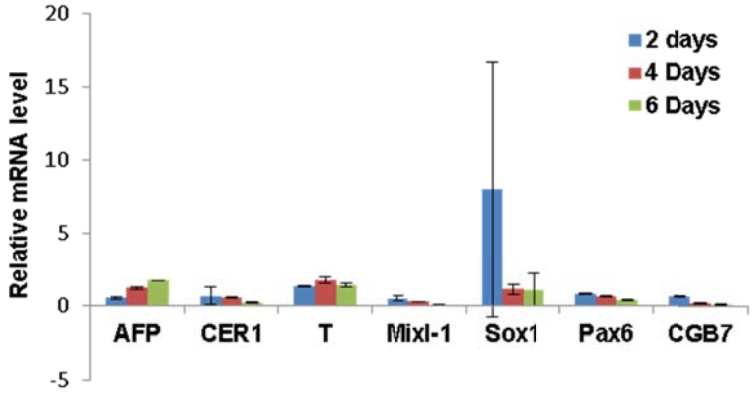


Figure 2.5. Displurigen does not induce lineage-specific differentiation. Quantitative-PCR analysis of endoderm (*AFP*, *CER1*), mesoderm (*T*, *MIXL1*), neuroectoderm (*SOX1*, *PAX6*), and trophectoderm (*CGB7*) markers in H1 hESCs treated with NSC375009 (Displurigen, 5 μM) for 2 - 6 days, compared to DMSO control. All values were normalized to the level (=1) of mRNA in cells without compound treatment. Each bar represents mean ± SD (error bars). *ACTB* (*β-actin*) was used as a loading control.

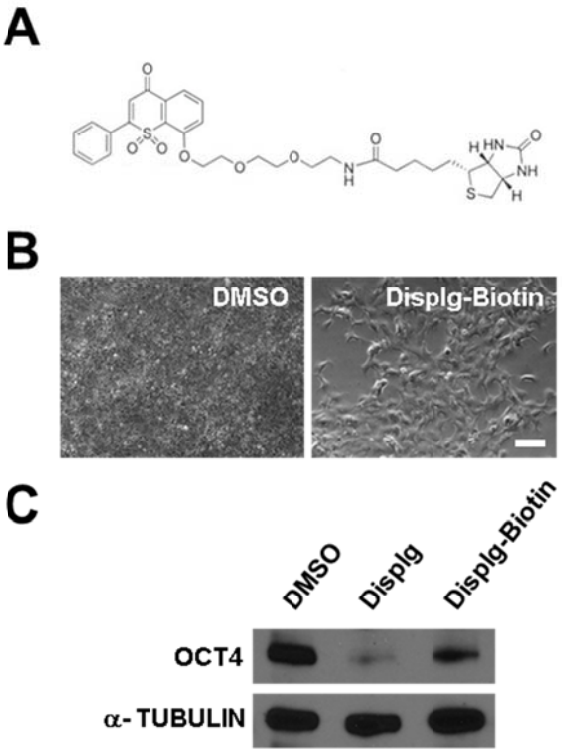


Figure 2.6. Biotinylated Displurigen disrupts hESC pluripotency. (**A**) Chemical structure of biotinylated Displurigen (Displg-Biotin). (**B**) Phase contrast images of H1 hESCs treated with DMSO and Displg-Biotin (100 μM) for 4 days. Scale bar: 100 μm. (**C**) Western blotting of OCT4 in H1 hESCs treated with DMSO, Displurigen (10 μM), and biotinylated Displurigen (Displg-Biotin, 100 μM) for 6 days. α-TUBULIN was used as a loading control.

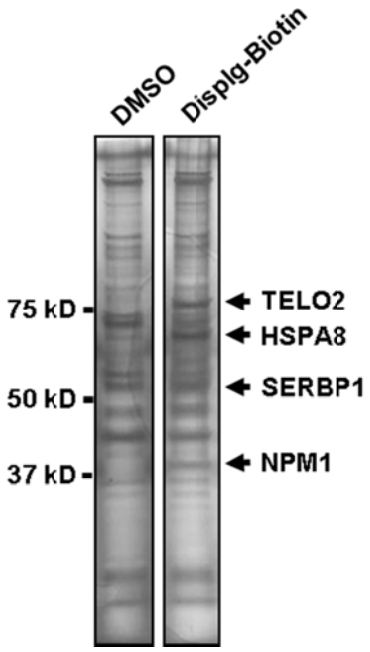


Figure 2.7. Biotinylated Displurigen pulls down HSPA8, TELO2, SERBP1, and NPM1. Silver staining of SDS-PAGE gel with proteins pulled down by Displg-Biotin. DMSO was used as a control. Mass spectrometry analysis revealed the Displg-Biotin-associated proteins as HSPA8-isoform-1, TELO2, SERBP1, and NPM1.

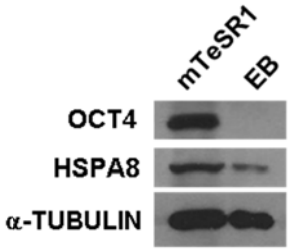


Figure 2.8. HSPA8 is expressed in undifferentiated hESCs and is downregulated upon differentiation. Western blotting of OCT4 and HSPA8-isoform-1 in undifferentiated H1 hESCs versus H1 hESCs subjected to embryoid body (EB) differentiation for 10 days. α-TUBULIN was used as a loading control.

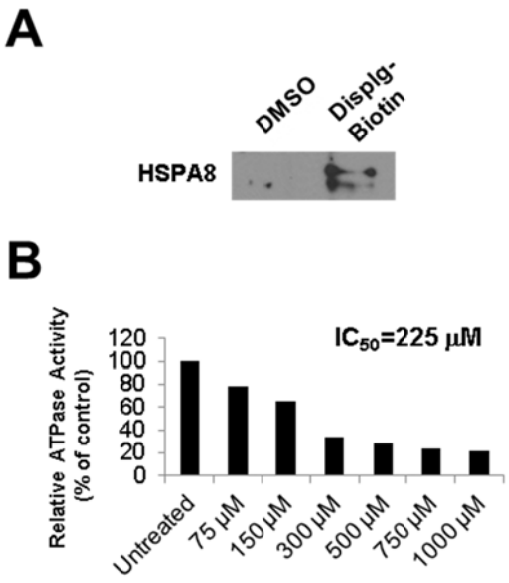


Figure 2.9. *In vitro* assays confirm HSPA8 as a target of Displurigen. (**A**) Western blotting of HSPA8-isoform-1 in the precipitates of Displg-Biotin (10 µM) and DMSO pull down assays, using H9 hESC lysates. Total lysates used were normalized between the DMSO and Displg-Biotin pull down assays. (**B**) *In vitro* ATPase activity of HSP70 in the presence of increasing doses of Displurigen. n = 1 independent experiment.

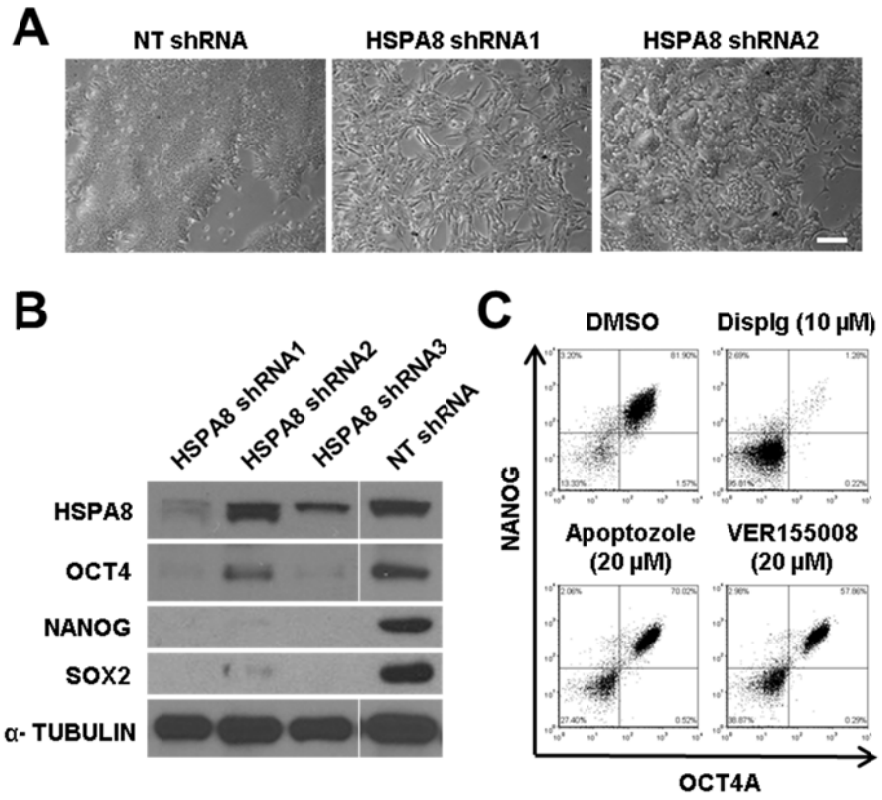


Figure 2.10. HSPA8 inhibition disrupts hESC pluripotency. (**A**) Phase-contrast images of H9 hESCs infected with lentivirus-particles containing Non-Target (NT) shRNA or shRNAs (shRNA1 and shRNA2) targeting HSPA8-isoform-1 for 5 days. Scale bar: 100 μm. (**B**) Western blotting of HSPA8-isoform-1, OCT4, SOX2, and NANOG in H1 hESCs infected with lentivirus-particles containing NT shRNA and HSPA8-1-targeting shRNAs. α-TUBULIN was used as a loading control. (**C**) FACS analysis of OCT4A and NANOG in H1 hESCs treated with DMSO, 10 μM Displurigen, 20 μM Apoptozole, and 20 μM VER155008 for 4 days.

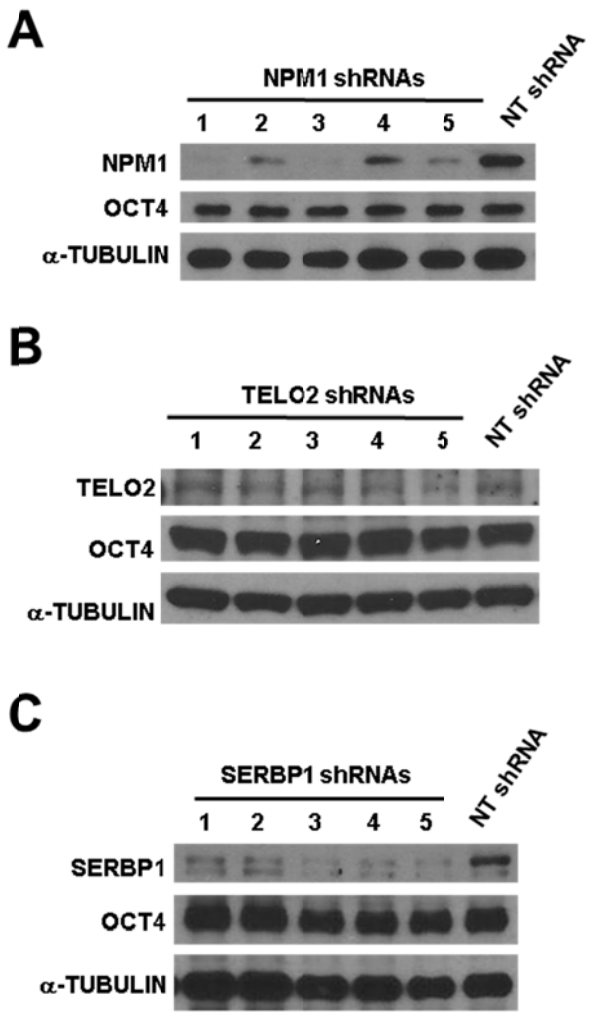


Figure 2.11. Knocking down NPM1, SERBP1, and TELO2 do not affect hESC pluripotency. (**A**) Western blotting of NPM1 and OCT4 in H1 hESCs infected by lentivirus-particles containing NT and NPM1 shRNAs for 7 days. α-TUBULIN was used as a loading control. (**B**) Western blotting of SERBP1 and OCT4 in H1 hESCs infected by lentivirus-particles containing NT and SERBP1 shRNAs for 4 days. α-TUBULIN was used as a loading control. (**C**) Western blotting of TELO2 and OCT4 in H1 hESCs infected by lentivirus-particles containing NT and TELO2 shRNAs for 13 days. α-TUBULIN was used as a loading control.

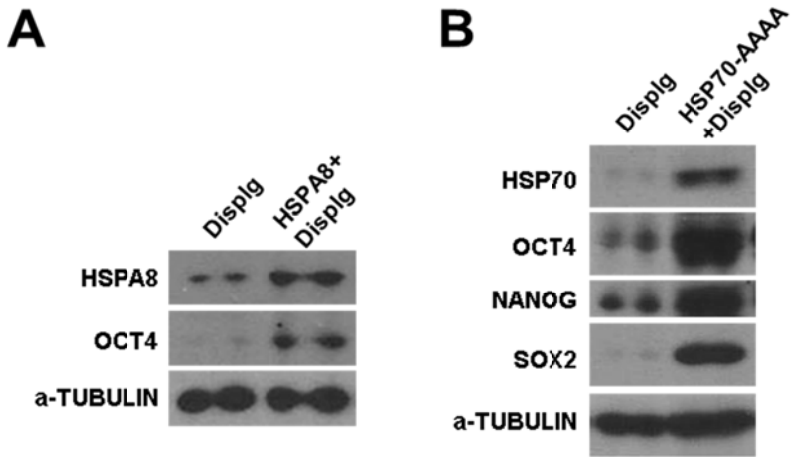


Figure 2.12. Overexpressing HSPA8 or HSP70-AAAA rescue the effects of Displurigen. (**A**) Western blotting of HSPA8-isoform-1 and OCT4 in H1 hESCs treated with 10 μM Displurigen (Displg) for 6 days with or without HSPA8 overexpression. α-TUBULIN was used as a loading control. (**B**) Western blotting of HSP70, OCT4, SOX2, and NANOG in H9 hESCs treated with 5 μM Displurigen (Displg) for 16 days with or without overexpression of HSP70-AAAA. α-TUBULIN was used as a loading control.

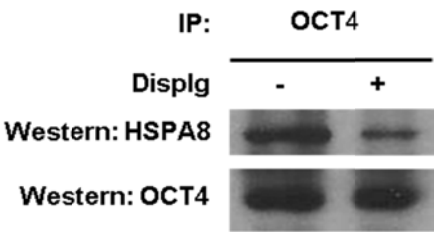


Figure 2.13. HSPA8 binds to OCT4 *in vivo*. Immunoprecipitation of OCT4 from H1 hESC lysates with or without Displurigen (1 mM), followed by Western blotting of OCT4 and HSPA8-isoform-1.

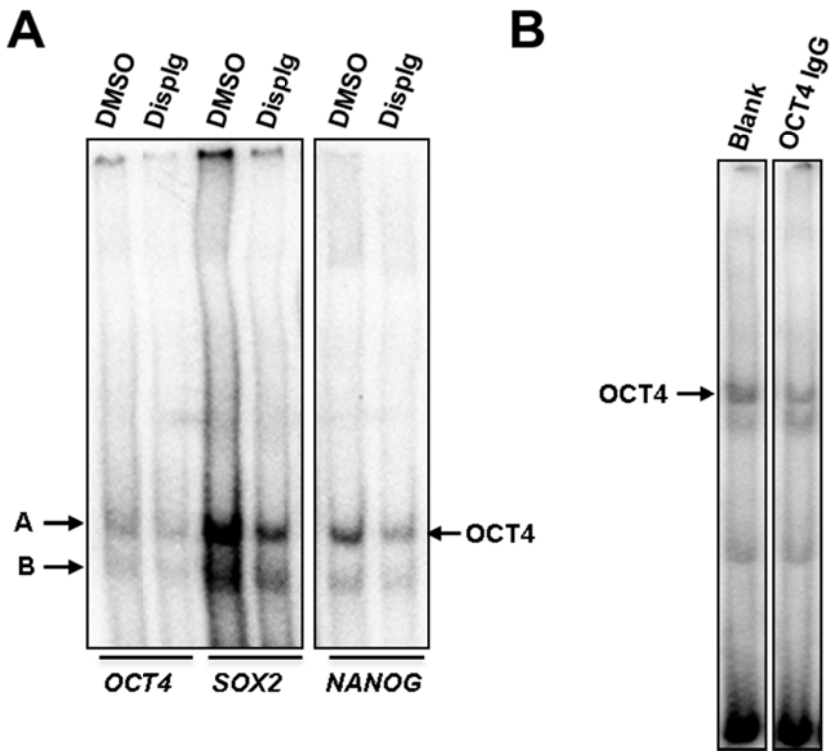


Figure 2.14. Displurigen inhibits the DNA-binding activity of OCT4. (**A**) EMSA using labeled OCT4-binding elements from *OCT4*, *SOX2*, and *NANOG* promoters to detect bindings of protein factors in the nuclear extracts of H9 hESCs with or without Displurigen treatment (10 μM, 24 hr). Labels A and B denote potential DNA-protein complexes, with band-A being the band representing OCT4-binding to the DNA probes. (**B**) EMSA using nuclear extracts of H9 hESCs treated with or without anti-OCT4 IgG showing the bindings of OCT4 proteins to the radio-labeled OCT4-binding element in *OCT4* promoter. The label "OCT4" points to the band sensitive to anti-OCT4 IgG treatment.

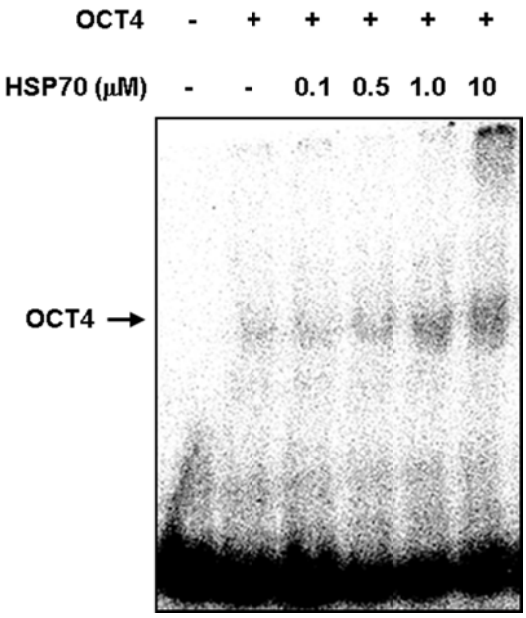


Figure 2.15. HSP70 directly facilitates the DNA-binding activity of OCT4 *in vitro*. EMSA showing binding (arrow) of purified recombinant OCT4 proteins (500 ng per reaction) to a labeled OCT4-binding element *in vitro*, with or without the addition of purified recombinant HSP70 proteins (up to 10 μM).

Tables

Antibody	Source	Catalog Number	Dilution
OCT4	Santa Cruz	sc-9081	1:1,000 (WB) / 1:100 (IF)
SOX2	Millipore	AB5603	1:500 (WB)
NANOG	Cell Signaling	4903P	1:500 (WB)
α-TUBULIN	Abcam	ab11304	1:10,000 (WB)
HSPA8-Isoform-I	R&D systems	MAB4148	1:100 (WB)
HSP70	Abcam	ab5439	1:1000 (WB)
TELO2	Santa Cruz	sc-138069	1:200 (WB)
SERBP1	Santa Cruz	sc-131741	1:200 (WB)
NPM1	Santa Cruz	sc-6013-R	1:200 (WB)
OCT4A-Alexa647	BD Biosciences	562252	1:20 (FACS)
NANOG-PE	BD Biosciences	560483	1:5 (FACS)

(WB: Western blotting; IF: immunofluorescence)

Table 2.1. Antibody sources and dilutions

Primer Name	Forward Primer (5'-3')	Reverse Primer (5'-3')
ACTB	agagctacgagctgcctgac	cgtggatgccacaggact
OCT4 (isoform A)	cttctcgcccctccaggt	aaatagaacccccagggtgagc
NANOG	tttggaagctgctggggaag	gatgggaggaggggagagga
SOX2	ggcagctacagcatgatgcaggagc	ctggtcatggagttgtactgcagg
T (Brachyury)	gctgtgacaggtacccaacc	catgcaggtgagttgtcagaa
MIXL1	ggcgtcagagtgggaaatcc	gcagttcacatctacctcaagag
CER1	acagtgcccttcagccagact	acaactacttttttcacagccttcgt
AFP	agcttggtggtggatgaaac	ccctcttcagcaaagcagac
CGB7	tccttggcgctagaccac	cagggaggcacaggagtg
SOX1	tacagccccatctccaactc	gctccgacttcaccagagag
PAX6	atttcccgctctggttcag	tagcgaagcctgacctctgt

Table 2.2. Names and sequences of qPCR primers

Target	shRNA	Sequence (5’ – 3’)
HSPA8-1	HSPA8-1-1	ccgggcaactgttgaagatgagaaactcgagtttctcatcttcaacagttgctttttg
	HSPA8-1-2	ccggccaagacttcttcaatggaaactcgagtttccattgaagaagtcttggttttttg
	HSPA8-1-3	ccgggcccgatttgaagaactgaatctcgagattcagttcttcaaatcgggctttttg
TELO2	TELO2-1	ccggatctgacctcacagttctatgctcgagcatagaactgtgaggtcagatttttttg
	TELO2-2	ccggagaacaagaaggcccagtttgctcgagcaaactgggccttcttgttctttttttg
	TELO2-3	ccggcctgatgtgcctggctgttaactcgagttaacagccaggcacatcaggttttttg
	TELO2-4	ccggtgaccacgtctgaggacatagctcgagctatgtcctcagacgtggtcatttttttg
	TELO2-5	ccggtagtgcagccagtccgctaaactcgagtttagcggactggctgcactatttttttg
SERBP1	SERBP1-1	ccgggcgcttaagaaagaaggaatactcgagtattccttctttcttaagcgcttttttg
	SERBP1-2	ccgggattcggttatggaccatcatctcgagatgatggtccataaccgaatctttttttg
	SERBP1-3	ccggcctgaaggtgaagaacatcatctcgagatgatgttcttcaccttcaggttttttg
	SERBP1-4	ccgggccgaggagatggatttgattctcgagaatcaaatccatctcctcggctttttttg
	SERBP1-5	ccggcctgaaggtgaagaacatcatctcgagatgatgttcttcaccttcaggttttttg
NPM1	NPM1-1	ccgggccaagaatgtgttgtccaaactcgagtttggacaacacattcttggctttttg
	NPM1-2	ccgggccgacaaagattatcactttctcgagaaagtgataatctttgtcggctttttg
	NPM1-3	ccgggcgccagtgaagaaatctatactcgagtatagatttcttcactggcgctttttg
	NPM1-4	ccgggcaaaggatgagttgcacattctcgagaatgtgcaactcatcctttgctttttg
	NPM1-5	ccggcctagttctgtagaagacattctcgagaatgtcttctacagaactaggttttttg

Table 2.3. Names and sequences of shRNAs

Probe Name	Forward Sequence (5'-3')	Reverse Sequence (5'-3')
NANOG	tctgcagctacttttgcattacaatggccttggtgag	ctcaccaaggccattgtaatgcaaaagtagctgcaga
OCT4	cagacagcagagagatgcatgacaaaggtgccgtgatggttc	gaaccatcacggcacctttgtcatgcatctctctgctgtctg
SOX2	gccgtttgccttcatttccataagaagattaaga	tcttaatcttcttatggaaatgaaggcaaacggc

Table 2.4. EMSA probes

CHAPTER 3: A NOVEL SMALL MOLECULE INHIBITOR OF TRPM6 PROMOTES NEARLY HOMOGENEOUS MESODERM AND DEFINITIVE ENDODERM DIFFERENTIATIONS OF HUMAN PLURIPOTENT STEM CELLS AND REVEALS A ROLE OF MAGNESIUM HOMEOSTASIS DURING GERM LAYER SPECIFICATION

Abstract

Cell-replacement therapies require differentiated cells of high purity, yet current human pluripotent stem cell-based differentiation protocols perform poorly in meeting this requirement. In this study, I optimized growth factor-directed mesoderm and definitive endoderm (DE) differentiation protocols to maximize their efficiencies. Moreover, through a chemical screen, I discovered a novel small molecule Mesendogen (MEG) which robustly induces nearly homogeneous (⩾85%) mesoderm and DE differentiations of human pluripotent stem cells when combined with the optimized protocols. Intracellular FACS analyses with stringent gate settings for multiple lineage-specific markers were used to rigorously discriminate and precisely quantify the mesoderm and DE progenitor populations generated in this study. Using a kinome screen I identified transient receptor potential cation channel, subfamily M, member 6 (TRPM6) as a potential target of MEG. The expression of TRPM6 is tightly controlled during animal development, and loss-of-function mutation of TRPM6 has been reported to cause developmental defects. The hetero-tetrameric channel complex formed by TRPM6 and TRPM7 is the major regulator of cellular and whole body magnesium homeostasis, yet the exact roles of TRPM6/TRPM7 channel and magnesium homeostasis during development are poorly understood. Loss-of-function experiments confirmed TRPM6 as the biological target of MEG. I further demonstrated that MEG reduces cellular magnesium level, and magnesium-withdrawal phenocopies MEG, indicating that MEG

may exert its biological function by inhibiting TRPM6/TRPM7 magnesium channel activity and thus modulating cellular magnesium uptake. This study describes a robust method for the highly dependable productions of nearly homogeneous mesoderm and DE progenitors, and reveals for the first time that TRPM6 and magnesium homeostasis are involved in the regulation of early germ layer specification.

Introduction

Fulfilling the promises of hPSC-based cell-replacement therapy requires the development of robust differentiation protocols that can produce terminally differentiated cellular progenies possessing relevant biological functions, in large quantities, and with high purities (Lee et al., 2009; Passier, 2003; Yabut and Bernstein, 2011). Because the majority of current differentiation protocols are built on a step-wise design (Cai et al., 2007; Chambers et al., 2009; Chen et al., 2009; D'Amour et al., 2006), each step of differentiation must be optimized in these protocols in order to achieve maximum efficiencies for the productions of the end products. For this purpose, bioactive small molecules have been widely applied to increase differentiation efficiencies toward multiple lineages with success (Borowiak et al., 2009b; Chen et al., 2009; Chen et al., 2012; Chetty et al., 2013; Gonzalez et al., 2011a; Lian et al., 2012; Mahmood et al., 2010; McLean et al., 2007; Zhu et al., 2009).

In the first part of this study, through a small scale chemical screen, I discovered a small molecule which I named Mesendogen (MEG for short) that robustly enhances the growth factor-induced mesoderm and DE differentiations of human embryonic stem cells

(hESCs) and human induced pluripotent stem cells (hiPSCs) to near-homogeneity (⩾ 85%). Both pre-treatment and combined treatment of MEG dramatically and consistently boosted the effects of mesoderm- and DE-inducing growth factors. The time window for this enhancement effect of MEG seems to be restricted to the very early stage of differentiation (within the first 24 hours), as delayed treatments of MEG failed to enhance differentiation.

Transient receptor potential cation channel, subfamily M, member 6 (TRPM6), together with another member of the TRPM protein family, transient receptor potential cation channel, subfamily M, member 7 (TRPM7), are a special type of kinase commonly referred to as the "channel kinases", in that they contain both a intracellular kinase domain and a transmembrane channel domain (Chubanov et al., 2005a; Chubanov et al., 2007; Chubanov et al., 2006; Chubanov et al., 2005b; Chubanov et al., 2004; Schlingmann et al., 2007). TRPM6 gene encodes 7 isoforms via alternative splicing, with isoforms TRPM6a, TRPM6b, and TRPM6c containing the transmembrane channel domain and the kinase domain, isoforms TRPM6-kinase-1, -2, and -3 containing only the kinase domain, and the testes-specific isoform TRPM6t containing only the transmembrane domain (Chubanov et al., 2004). Hetero-tetramer channel complexes formed between TRPM6 and TRPM7 and homo-tetrameric TRPM7 channels are active and are responsible for magnesium ion transportation (Chubanov et al., 2005b; Chubanov et al., 2004; Schlingmann et al., 2007), while homo-tetramer channels formed by TRPM6 alone are not functional (Zhang et al., 2014).

TRPM7 is ubiquitously expressed in all tissue types, whereas the expression of TRPM6 is regulated during development and is eventually confined to a few tissue types (Chubanov et al., 2005a; Walder et al., 2009). TRPM6 has been reported to play a role during embryonic development and in embryonic stem cell maintenance: TRPM6-knockout in mouse caused embryonic mortality and neural tube defects (Walder et al., 2009); homozygous TRPM7-deficient mice were also embryonic lethal, while mouse embryonic stem cells deficient in TRPM7 displayed growth arrest which could be rescued by Mg^{2+} supplementation (Ryazanova et al., 2010).

In the second part of this study I report that using target identification techniques followed by functional validations, I identified TRPM6 as the biological target of MEG. Using a magnesium sensor I demonstrated that MEG effectively reduces cellular magnesium level. I also showed that both inhibition of the magnesium-importing activity of TRPM6/TRPM7 channel complex by known channel blockers and withdrawal of Mg^{2+} in the differentiation medium gave rise to similar phenotypes as compared to MEG treatment. These data suggest that MEG may exert its biological function by acting as an inhibitor of TRPM6/TRPM7 channel activity and, subsequently, by modulating cellular Mg^{2+} uptake. This study uncovers a novel molecular mechanism which governs early mesoderm and DE inductions through regulation of magnesium homeostasis.

Results

Small scale chemical screening identified a novel small molecule that induces mesoderm and endoderm, but not neural differentiation of hESCs

In Chapter 2 I described the identification of 29 bioactive small molecules that potently disrupt hESC pluripotency. Using this compound collection I conducted a screen (Materials and Methods) to search for small molecules that specifically induce mesoderm and DE differentiations, and identified compound 6528694 (N-({[2-chloro-5-(trifluoromethyl)phenyl]amino}carbonothioyl)-4-isopropylbenzamide; Figure 3.1A). A 7-day treatment of 6528694 in human embryonic stem cells (hESCs) maintained under a pluripotency culture condition ("mTeSR1" condition hereafter) (Ludwig et al., 2006) induced elevated protein expressions of mesoderm markers T (Brachuary) and EOMES (Eomesodermin) and DE markers SOX17 and FOXA2, but not neural-specific markers PAX6 and SOX1 (Figure 3.1B). Meanwhile pluripotency was disrupted as expected, shown by downregulations of the pluripotency markers OCT4 and SOX2 (Figure 3.1B). To further confirm its selectivity against the neural lineage, I incubated 6528694 with hESCs undergoing neural differentiation driven by a BMP pathway inhibitor Dorsomorphin (Kim et al., 2010; Mak et al., 2012; Zhou et al., 2010b) (Materials and Methods). Addition of 6528694 completely abolished the neural induction effect of Dorsomorphin, as demonstrated by the failure of induction for SOX1 and PAX6 (Figure 3.1C). I thus name this compound Mesendogen, short for Mesoderm and Definitive Endoderm Inducing Reagent, and hereafter refer to this molecule as MEG for short.

The sole fact that lineage-specific markers were elevated upon MEG treatment cannot prove that mesoderm or DE progenitors have been successfully generated. To

prove the existence of such a population, one must demonstrate the emergence of a distinct cell population that expresses progenitor markers at the single cell level. However, MEG treatment alone was insufficient to induce such a distinct mesoderm progenitor population: when incubated with hESCs in a growth factor-free mesoderm basal differentiation medium (Materials and Methods), MEG marginally increased T expression in the overall population, but failed to drive the differentiation of a distinct T^+ population compared to the DMSO control, as shown by FACS analysis (Figure 3.2). I thus decided to test whether MEG could enhance the efficiencies of growth factor-guided mesoderm and DE differentiation protocols at generating mesoderm and DE progenitors.

MEG induces nearly homogeneous mesoderm differentiation in combination with growth factors

To test the effect of MEG in growth factor-induced mesoderm differentiation, I first derived a 2-dimensional (2-D) growth factor-guided mesoderm induction protocol (Figure 3.3 and Materials and Methods; "A-BVF" method hereafter) modified based on previous publications (Bernardo et al., 2011; Evseenko et al., 2010; Yu et al., 2011). Briefly, Activin A (A), bone morphogenetic protein 4 (BMP4; B), vascular endothelial growth factor (VEGF; V), and basic firoblast growth factor (bFGF; F) were used to drive differentiation; cells were analyzed after 1.5 - 2 days of induction. Using this method, I found that mesoderm differentiation efficiencies highly depended on the initial plating densities of hESCs: the lowest plating density tested (0.1×10^5 cells/cm^2) gave the highest yield of T^+EOMES^+ cells (73.1±3.2% by FACS; Figure 3.4) which represent mesoderm progenitors; this efficiency quickly dropped to near-zero at and above 0.5×10^5 cells/cm^2 (Figure 3.4).

I then added MEG to the A-BVF induction protocol, either by directly combining with the growth factors (Figure 3.3; "MEG+A-BVF" or "combination" condition hereafter) or by pre-incubating in mTeSR1 medium for 24 hours before starting the growth factor treatments (Figure 3.3; "MEG->A-BVF" or "pre-treatment" condition hereafter). Both approaches dramatically improved the induction efficiencies of A-BVF in all plating densities tested (Figure 3.4). Most notably, the combination condition enhanced the differentiation efficiency (measured by percentage of T^+/$EOMES^+$ cells) at the density of 0.1×10^5 cells/cm^2 from 73.1±3.2% for the A-BVF condition to an average of 82.7±3.6%, whereas the pre-treatment condition dramatically enhanced the efficiency at 0.25×10^5 cells/cm^2 from 25.8±10.3% to 86.1±1.3% (Figure 3.4). An example of nearly homogeneous mesoderm differentiation at 0.25×10^5 cells/cm^2 plating density under both the pre-treatment condition and the combination condition is shown in Figure 3.5. Similar results were acquired through immunofluorescence staining (Figure 3.6) and Western blotting (Figure 3.7) using the same markers T and EOMES, and through qPCR using a range of mesoderm markers including *T*, *EOMES*, *MIXL1*, *EVX1*, *TBX6*, *HAND1*, *MESP1*, and *MEOX1* (Figure 3.8).

MEG induces nearly homogeneous definitive endoderm differentiation in combination with growth factors

To test whether MEG enhances growth factor-induced DE differentiation in a similar fashion as compared to its effects in mesoderm differentiation, I first established a 2-D growth factor-guided DE induction protocol (Figure 3.9 and Materials and Methods; "AWS" method hereafter) designed based on reported methods (Borowiak et al., 2009b;

D'Amour et al., 2005) with modifications. Briefly, Activin A (A), WNT3A (W), and increasing concentrations of fetal bovine serum (S) were applied to drive differentiation; cells were analyzed after 5 - 7 days of induction. Using this method I found that, similar to mesoderm differentiation, DE differentiation efficiencies also highly depended upon the initial plating densities, as lower densities again gave rise to higher differentiation efficiencies as measured by percentage of $SOX17^{+}FOXA2^{+}$ cells which represent DE progenitors, using FACS analysis (Figure 3.10). The highest differentiation efficiencies induced by the AWS condition were given by cells plated at 0.25×10^5 cells/cm^2 (40.6±30.2%; Figure 3.10).

I then added MEG to the AWS condition either by pre-treatment (Figure 3.9; "MEG->AWS" or "pre-treatment" condition) or in combination with the growth factors (Figure 3.9; "MEG+AWS" or "combination" condition). Both approaches significantly boosted differentiation efficiencies for all plating densities tested (Figure 3.10). Most notably, both treatments dramatically enhanced the differentiation efficiency at the density of 1.0×10^5 cells/cm^2, from 11.8±8.7% $SOX17^{+}FOXA2^{+}$ for the AWS condition, to an average of 72.7±19.7% for the combination condition and 87.7±7.6% for the pre-treatment condition (Figure 3.10). In some experiments, MEG treatments dramatically enhanced the efficiencies of DE progenitor differentiation to >90% under this optimum plating density, as shown by FACS analysis (Figure 3.11). A time course analysis of the MEG->AWS condition by immunofluorescence staining of SOX17 and FOXA2 is shown in Figure 3.12, also demonstrating a nearly homogeneous $SOX17^{+}FOXA2^{+}$ culture on day 7. The expressions of several other endoderm-specific markers including CER1,

GATA4, GATA6, and GSC were also significantly enhanced by MEG, as shown by Western blotting (Figure 3.13) and qPCR (Figure 3.14). It should be noted that SOX17, FOXA2, and all the other endoderm markers discussed so far are pan-endoderm markers that are expressed in both definitive and extra-embryonic endoderm lineages (Borowiak et al., 2009b). In order to rule out the possibility of extra-embryonic endoderm (ExEn) differentiation in this method, I examined the expressions of ExEn-specific markers AFP (α-fetoprotein) and SOX7 using Western blotting (Figure 3.13) and qPCR (Figure 3.14). Results showed no significant elevation of AFP and SOX7 in all differentiation conditions examined (Figures 3.13 and 3.14), verifying the DE-lineage identity of the differentiated progenies.

I find one technical question particularly worth noting for the FACS analyses conducted in this study, which is that undifferentiated hESCs already possess significant background signals for T, EOMES, SOX17, and FOXA2 stainings before undergoing differentiations. Therefore, if isotype controls were used to set up quadrants for the dot plots, these background signals will lead to extremely high false positive rates for both mesoderm (98.42% T^+EOMES^+; Figure 3.15) and DE (46.74% $SOX17^+FOXA2^+$; Figure 3.16) differentiations. Thus, undifferentiated hESCs stained with T and EOMES antibodies (for mesoderm differentiation experiments) or with SOX17 and FOXA2 antibodies (for definitive endoderm differentiation experiments), instead of isotype IgG controls, were used as gating controls for all FACS analyses performed in this study to ensure a stringent measurement of the differentiation efficiencies.

MEG possesses robust differentiation-enhancing activities to human induced pluripotent stem cells (hiPSCs) and to hESCs under various mesoderm and DE differentiation conditions

The differentiation-enhancing effects of MEG are robust and are not limited to hESCs or to the above mentioned A-BVF and AWS differentiation conditions. The first point is evidenced by the fact that MEG effectively enhanced the mesoderm differentiation efficiency of a human induced pluripotent stem cell (hiPSC) line MMW2 (Mali et al., 2010) in a similar fashion as compared to its effect in hESCs (Figure 3.17). A DE differentiation-enhancing effect of MEG similar to what was found in hESCs was also observed in the hiPSC line MMW2 (Figure 3.18).

MEG also improved the efficiencies of several variations of mesoderm and DE induction conditions adopted from reported methods with modifications, including BVF (Bernardo et al., 2011) (BMP4, VEGF, and bFGF; Figure 3.19A) and BF (Bernardo et al., 2011; Yu et al., 2011) (BMP4 and bFGF; Figure 3.19B) conditions for mesoderm differentiation, and AW-B27 (Norrman et al., 2013) (serum replaced by B-27 supplement; Figure 3.20A), AW-ITS (Cai et al., 2007) (serum replaced by ITS supplement; Figure 3.20B), and AS (Borowiak et al., 2009b) (withdrawal of WNT3A; Figure 3.20C) conditions for DE differentiation, demonstrating that MEG possesses general mesoderm and DE differentiation-enhancing activities that are not confined to the A-BVF and AWS conditions used in this study. Detailed protocols for these variant differentiation conditions are described in Materials and Methods.

Delayed treatments of MEG (initiate on day 1 rather than on day 0) were ineffective

Based on the results that the pre-treatment conditions of MEG in both mesoderm and DE differentiations gave rise to similar, if not better, differentiation efficiencies compared to the combination conditions, I hypothesized that MEG may function by triggering a pathway (or pathways) involved only in the very early phases of mesoderm and DE inductions. Indeed, for both MEG+A-BVF and MEG+AWS conditions, the differentiation-enhancing effects of MEG were nearly completely lost when MEG was added starting from day 1 of differentiation instead of from day 0 (Figures 3.21A and 3.21B), strongly supporting this hypothesis.

In vitro kinome screen and functional validations identified TRPM6 as the biological target of MEG

To investigate the molecular mechanism of MEG, I employed target identification techniques (Schenone et al., 2013; Ziegler et al., 2013). The enzymatic activities of kinases are typically inhibited at their ATP-binding sites by small planar compounds with no stereocenters and high aromatic contents (Huigens et al., 2013). Due to the planar structure and the lack of stereocenters of the MEG molecule, I postulated that MEG might function by inhibiting proteins containing ATP-binding sites, such as kinases. I thus applied an *in vitro* kinase screen using a commercial service named KINOMEscan (DiscoveRx; Materials and Methods). A total of 456 kinases were analyzed, with results summarized in Figure 3.22A: lower Percent-of-Control (%Ctrl) values indicate stronger interactions between the test compound and the kinases. Only two kinases gave %Ctrl values lower than 60, with %Ctrl values of 50 for TRPM6 and 47 for NEK3 (Figure 3.22B).

A previous study showed that TRPM6-knockout mouse demonstrated developmental defects in neural tube formation (Walder et al., 2009), indicating a potential involvement of TRPM6 in the regulation of early embryonic development. I thus hypothesized that TRPM6 may be the biological target of MEG.

To test this hypothesis I first confirmed the expression of TRPM6 in undifferentiated hESCs. Indeed, RT-PCR analysis using TRPM6 isoform-specific primers (Chubanov et al., 2004) showed that the kinase- and channel domain-containing isoforms TRPM6a and TRPM6b, but not the other isoforms of TRPM6, were expressed in undifferentiated hESCs (Figure 3.23). TRPM7, the ubiquitously expressed member of the TRPM protein family (Chubanov et al., 2005a), was also expressed in hESCs as expected (Figure 3.23).

I then tested whether the inhibition of TRPM6 expression can mimic the effect of MEG on hESC differentiation. Indeed, TRPM6 knockdown significantly enhanced growth factor-induced mesoderm differentiation to an extent comparable to MEG-induction, as measured by the percentage of $T^{+}EOMES^{+}$ mesoderm progenitor cells in differentiated cultures (Figure 3.24). TRPM6 knockdown also gave similar enhancement effects to the growth factor-induced DE differentiation with the efficiencies measured by percentage of $SOX17^{+}FOXA2^{+}$ DE progenitor cells (Figure 3.25). On the other hand, knockdown of NEK3 failed to offer any consistent enhancement effect during growth factor-induced mesoderm (Figure 3.26A) and DE differentiations (Figure 3.26B). The

shRNA constructs of TRPM6 and NEK3 were delivered to cells via transient transfection at the onset of differentiation in order to avoid the possibility that stable knockdown by lentiviral delivery may affect the survivability of the infected cells, leading to the selective loss of infected cells by the end of a differentiation cycle and thus inaccurate measurement of the differentiation efficiency (Materials and Methods). These results, from a phenotypical standpoint, confirmed that TRPM6, but not NEK3, is more likely to be the biological target of MEG.

Inhibitions of TRPM6/TRPM7 channel activity and cellular Mg^{2+} uptake mimic the biological effect of MEG

As the best understood biological function of TRPM6 is in magnesium ion transportation (Chubanov et al., 2005b; Chubanov et al., 2004; Schlingmann et al., 2007), I set out to investigate the effect of MEG on cellular magnesium homeostasis. Indeed, overnight incubation of MEG dramatically decreased intracellular magnesium level in a dose dependent manner, as shown by flow cytometry analysis using a magnesium sensor Mag-fluo-4 as indicator (Figure 3.27). I thus hypothesized that MEG functions by inhibiting the magnesium transport activity of TRPM6/TRPM7 channel, and tested this hypothesis by examining whether other known TRPM6/TRPM7 channel blockers or Mg^{2+}-withdrawal can phenocopy MEG.

The addition of TRPM6/TRPM7 channel blockers Rethenium Red (50 μM) and 2-APB (50 μM) indeed significantly enhanced mesoderm differentiation efficiency to the same degree as observed in MEG induction (10 μM) (Figure 3.28); similar results were obtained in DE differentiation with 2-APB at 50 μM and Rethenium Red at 10 μM

(Figure 3.29). Meanwhile, withdrawal of Mg^{2+} in mesoderm differentiation medium dramatically enhanced growth factor-induced mesoderm differentiation to a degree even greater than that achieved by MEG treatment (74.3% versus 42.1% T^+EOMES^+; Figure 3.30), whereas the combined treatment of MEG failed to further enhance mesoderm differentiation efficiency on top of Mg^{2+}-withdrawal (73.4% versus 74.3% T^+EOMES^+; Figure 3.30), indicating that withdrawal of Mg^{2+} not only phenocopied MEG but also rendered MEG-treatment ineffective for further mesodermal induction. Similar results were observed in DE differentiation, with Mg^{2+}-withdrawal during the first day of differentiation significantly boosted the overall efficiency from 49.8% to 75.9% $SOX17^+FOXA2^+$ (Figure 3.31). Taken together, these results demonstrate that MEG most likely exerts its biological function by inhibiting the TRPM6/TRPM7 channel activity and, subsequently, cellular magnesium homeostasis.

Discussion

In summary, this study demonstrates nearly homogeneous derivations of mesoderm and DE progenitors from hPSCs enhanced by a novel small molecule inhibitor MEG. Further investigation indicates that MEG performs its biological function by targeting TRPM6 and subsequently regulating cellular magnesium level.

Extensive efforts were taken to scrutinize the identities of the differentiated progenies and to ensure accurate measurements of the differentiation efficiencies in this study. This includes the applications of multiple experimental methods such as FACS analysis, immunofluorescence staining, Western blotting, and quantitative PCR. Most

notably, the method of intracellular FACS staining of lineage-specific markers, arguably one of the most accurate and discriminative method currently available for cell-fate validation at the single cell level, was used as the primary method in this study to distinguish and quantify the productions of mesoderm and DE progenitors.

The fact that MEG, TRPM6/TRPM7 channel blockers, and magnesium-withdrawal simultaneously enhanced the inductions of both mesoderm and DE, the two daughter lineages of primitive streak (or mesendoderm), indicates that a low intracellular level of magnesium may be favorable to primitive streak formation during early human development. The fact that delayed treatment of MEG was ineffective also indicates that this biological phenomenon may be most relevant at the earliest onset of primitive streak induction *in vivo*.

In summary, this study develops a robust and easily applied chemical tool for directed differentiation which could easily be routinely integrated into existing hPSC-based differentiation protocols that aim at deriving mesoderm and DE lineage progenies. This study also identified TRPM6 as the endogenous protein target of the small molecule MEG, which in turn shed light to a potentially novel function of magnesium during early human development, especially germ layer specification. Last but not least, as a potential chemical inhibitor of TRPM6/TRPM7 channel complex, the discovery of MEG could provide the scientific community with a novel tool for the basic research of TRPM6/TRPM7 channel complex or a novel drug candidate for diseases associated with

abnormal TRPM6/TRPM7 channel activities and magnesium homeostasis (Nair et al., 2012; Swaminathan, 2003; Wolf and Trapani, 2012).

Materials and Methods

Cell culture

H9 and H1 hESC lines (WiCell Research Institute, Madison, WI) and MMW2 hiPSCs (Mali et al., 2010) were maintained under a feeder condition or a feeder-independent condition. For the feeder condition (Thomson et al., 1998), primary mouse embryonic fibroblasts (MEFs) prepared from embryos of pregnant CF-1 mice (day 13.5 of gestation; Charles River) were cultured in DMEM containing 10% FBS (Hyclone), 1% non-essential amino acids (NEAA; Invitrogen), and penicillin/streptomycin, and mitotically inactivated by gamma irradiation. H9 and H1 hESCs and hiPSCs were cultured on irradiated MEFs in media containing DMEM/F12, 20% knockout serum replacement (KSR; Invitrogen), 4 ng/ml basic fibroblast growth factor (bFGF; Invitrogen), 1% NEAA, 1 mM glutamine, and 0.1 mM β-mercaptoethanol. For the feeder-independent condition, hESCs and hiPSCs were cultured on Matrigel (BD Biosciences)-coated plates in mTeSR1 medium (StemCell Technologies) as described (Ludwig et al., 2006). Experiments described in this study were conducted with H9 and H1 hESCs between passages 30 and 60.

Small scale screening of mesoderm and DE inducers

For this screen, hESCs were maintained in the pluripotent state in mTeSR1 medium (Ludwig et al., 2006). Small molecules were added and incubated for 5 - 7 days

to induce differentiation. Differentiated cultures were then assayed for mesoderm, DE, and neural lineage-specific marker expressions using Western blotting. Compounds that selectively induced the expressions of mesoderm and DE markers, but not neural markers, were identified as hits.

Neural differentiation

For neural differentiation hESCs were dissociated by Accutase (Invitrogen) and seeded onto culture plates pre-coated with Matrigel (BD Biosciences) in mTeSR1 medium (StemCell Technologies) supplemented with Y-27632 (10 μM; Calbiochem). Cultures were let grow till complete confluency and then changed into neural induction medium composed of Advanced DMEM/F12, 1% N-2 supplement, and Glutamax (all from Invitrogen) and supplemented with or without Dorsomorphin (1 μM; Sigma). Medium was changed every day. Cells were collected on day 8 for analysis.

Mesoderm differentiation

To induce directed mesoderm differentiation, we used a growth-factor combination ("A-BVF"), as previously reported with modifications (Bernardo et al., 2011; Evseenko et al., 2010; Yu et al., 2011). H1 and H9 hESCs or MMW2 hiPSCs were dissociated by Accutase (Invitrogen) and seeded onto culture plates pre-coated with Matrigel (BD Biosciences) in mTeSR1 medium (StemCell Technologies) supplemented with Y-27632 (10 μM; Calbiochem). The day of seeding was referred to as "day -2". On day 0, mTeSR1 medium was replaced by a mesoderm induction medium, in which the basal medium (Advanced RPMI-1640 basal medium supplemented with B-27 and

Glutamax; all from Invitrogen) was supplemented by Activin A (R&D Systems), BMP4 (Invitrogen), human VEGF (Invitrogen), and bFGF (Invitrogen) (all at 10 ng/ml) from day 0 to day 1, while Activin A was removed from this medium from day 1 to day 2. This protocol was referred to as the A-BVF condition.

For MEG induction, MEG was either added to the mTeSR1 medium on day "-1" as the "MEG->A-BVF" or "pre-treatment" condition (10 μM), or added to the mesoderm induction medium as the "MEG+A-BVF" or "combination" condition (1, 5, or 10 μM). Cells were collected on day 1.5 to day 2 of this protocol for further analyses.

Variations of mesoderm induction protocols used in this study include the "BVF" condition with BMP4 (10 ng/ml), VEGF (10 ng/ml), and bFGF (20 ng/ml), and the "BF" condition with BMP4 (10 ng/ml) and bFGF (20 ng/ml).

Definitive endoderm differentiation

To induce definitive endoderm differentiation, we applied a method utilizing Activin A, WNT3A and low concentrations of fetal bovine serum (FBS) which was modified from previously reported methods (Borowiak et al., 2009b; D'Amour et al., 2005). H1 and H9 hESCs or MMW2 hiPSCs were seeded in the same way as described for mesoderm differentiation, on day "-2". On day 0, mTeSR1 medium was replaced by an definitive endoderm induction medium, in which a basal medium (Advanced RPMI-1640 medium supplemented by Glutamax; Invitrogen) was supplemented by Activin A (50 or 100 ng/ml; from day 0 to day 7), WNT3A (25 ng/ml; from day 0 to day 1) (R&D

Systems), 0.2% FBS (from day 1 to day 2), and 2% FBS (from day 2 to day 7). This protocol was referred to as the AWS condition.

For MEG induction, MEG was either added to the mTeSR1 medium at 10 μM on day "-1" as the "MEG->AWS" or "pre-treatment" condition, or added to the definitive endoderm induction medium at 1 μM throughout the course of differentiation as the "MEG+AWS" or "combination" condition. Cells were collected on day 5 of this protocol for qPCR analysis, and day 7 for Western blotting or FACS analysis.

Variations of DE induction protocols used in this study include the "AW-B27" condition in which FBS was replaced by 2% B-27 supplement (Invitrogen), the "AW-ITS" condition in which FBS was replaced by Insulin-Transferrin-Selenium supplement (ITS; Invitrogen) at concentrations of 0.1% from day 1 to day 2 and 1% from day 2 to day 7, and the "AS" condition in which WNT3A was removed from the growth factor combination.

Flow cytometry

Single cell suspensions were acquired through Accutase (Invitrogen) treatment. Cells were fixed and stained using the Transcription Factor Buffer Set (BD Biosciences) following the manufacturer's instructions. Conjugated antibodies were used, including T-PE, EOMES-APC, FOXA2-Alexa488, and SOX17-APC (all from R&D Systems). Cells were resuspended in PBS supplemented with 1% BSA and analyzed using a BD Biosciences LSR II flow cytometry analyzer and BD FACSDiva software.

Undifferentiated cells cultured in mTeSR1 medium and stained with T and EOMES antibodies (for mesoderm) or SOX17 and FOXA2 antibodies (for definitive endoderm), instead of isotype IgG controls, were used as gating controls. This is due to the fact that undifferentiated hESCs already possess significant background signals for T, EOMES, SOX17, and FOXA2 staining before undergoing mesoderm or DE differentiations. If isotype controls were used to set up quadrants, these background signals will lead to extremely high false positive rates for both mesoderm (98.42% T^+EOMES^+; Figure 3.15) and DE (46.74% $SOX17^+FOXA2^+$; Figure 3.16) differentiations.

Western blotting

Cultured cells were lysed directly by 2× Laemmli buffer (Bio-Rad), boiled for 5 min, and analyzed using SDS-PAGE electrophoresis followed by wet-transfer onto nitrocellulose membranes using a system manufactured by Bio-Rad. The membranes were blocked using blocking solution (5% BSA in Tris-buffered saline containing 0.1% Tween-20 [TBST]), and then incubated with primary antibodies, diluted in TBST, at 4°C overnight. The membranes were then washed by TBST for 3 × 5 min, and incubated with horse radish peroxidase (HRP) conjugated secondary antibodies at room temperature for 1 hr. Finally, the membranes were washed 3 - 5 × 5 min by TBST and developed using Super-Signal West Pico Chemiluminescent Substrate (Pierce).

Immunofluorescent staining

Cells were washed once with phosphate-buffered saline (PBS) and fixed by 4% paraformaldahyde (Santa Cruz) at room temperature for 20 min, permeabilized by 0.5% Triton X-100 (Sigma) in PBS (PBST) for 15 min, and then blocked with 5% donkey serum (Sigma) in 0.1% PBST at room temperature for 1 hr. The samples were incubated with primary antibodies in 0.1% PBST at 4°C overnight, washed three times by PBS, and then incubated with fluorescent-labeled secondary antibodies in 0.1% PBST at room temperature for 2 hr. Finally, the cells were incubated with DAPI for 5 min, washed three times by PBS, and subjected to fluorescent microscopy analysis (Zeiss).

RNA extraction, reverse transcription, and quantitative-PCR

Total RNA were isolated using RNeasy mini kit (QIAGEN). cDNAs were synthesized from the purified RNAs using Reverse Transcription System (Promega). Quantitative-PCR was performed using QuantiTech SYBR Green PCR kit (QIAGEN). Signals were analyzed using the comparative C_T method, and ACTB gene was used as an internal control.

For TRPM6 isoform detection, cDNAs extracted from undifferentiated hESCs were used as templates for PCR amplification of the isoforms followed by agarose gel electrophoresis. For the sequences of isoform-specific PCR primers see Chubanov et al (Chubanov et al., 2004).

KINOMEscan (DiscoveRx)

In this screen, DNA-tagged recombinant kinases were bound by ligands linked to a solid support. When a small molecule inhibitor was added, it will bind to its target kinase(s), either at the ligand-binding sites which directly competes with the ligand, or at other parts of the protein which indirectly inhibits ligand-binding of the kinase by inducing conformational changes. This inhibition will release the target kinase of this small molecule from its ligand, and kinases released from the solid support will be detected by PCR analysis through their DNA tags. Results for binding interactions were reported as Percent of Control (%Ctrl): %Ctrl = [(positive control signal – test compound signal) / (positive control signal – negative control signal)] × 100. DMSO was used as negative control. Lower values of %Ctrl indicate stronger interactions between the test compound and the kinases.

Fugene HD transfection

Fugene HD reagent was purchased from Promega. Transfections were conducted following manufacturer's instructions. Briefly, plasmid DNA and Fugene HD transfection reagent were mixed at a 3:1 ratio in Opti-MEM medium (Invitrogen), incubated 15 min at room temperature, and applied to cells at day -1 or day 0 of the A-BVF or AWS differentiation protocols. ROCK inhibitor Y-27632 was added during transfection to boost transfection efficiency as previously reported (Yen et al., 2014).

Intracellular magnesium analysis

The cell permeant magnesium ion sensor Mag-fluo-4 was purchased from Invitrogen. 10 μM Mag-fluo-4 was loaded into cells for 30 min at room temperature in the dark, washed with PBS, and then analyzed using a BD Biosciences LSR II flow cytometry analyzer and BD FACSDiva software.

Statistical analysis

Statistical analyses were performed using Microsoft Excel.

Figures

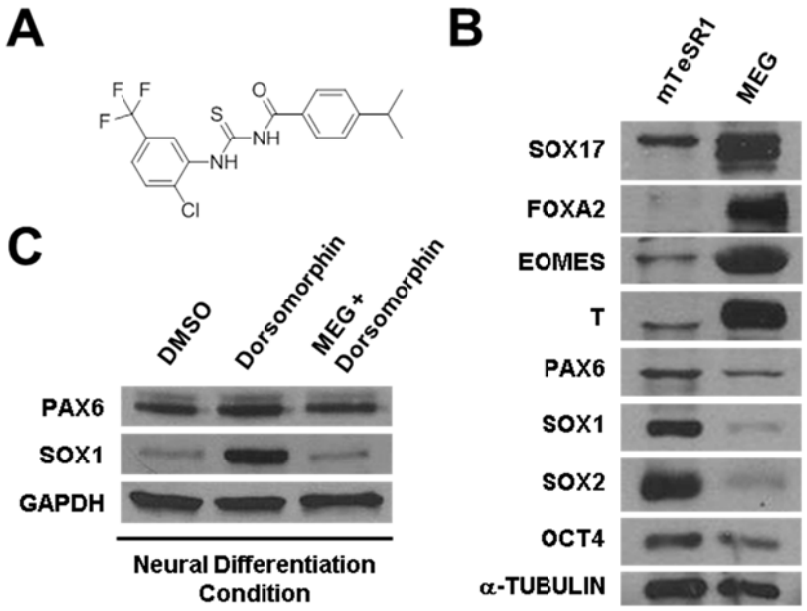


Figure 3.1. Compound 6528694 (Mesendogen) induced mesoderm and endoderm differentiation, but inhibits neural differentiation. (**A**) Chemical structure of compound 6528694 (Mesendogen or MEG). (**B**) Western blotting of endoderm markers SOX17 and FOXA2, mesoderm markers T and EOMES, neural-specific markers SOX1 and PAX6, and pluripotency markers OCT4 and SOX2 in H9 hESCs untreated (mTeSR1) or treated with 6528694 (MEG, 10 μM) for 7 days while maintained in mTeSR1 medium. α-TUBULIN was used as a loading control. (**C**) Western blotting of neural progenitor markers SOX1 and PAX6 in H1 hESCs treated with DMSO, Dorsomorphin (1 μM) and 6528694 (MEG, 10 μM) plus Dorsomorphin for 8 days. GAPDH was used as a loading control.

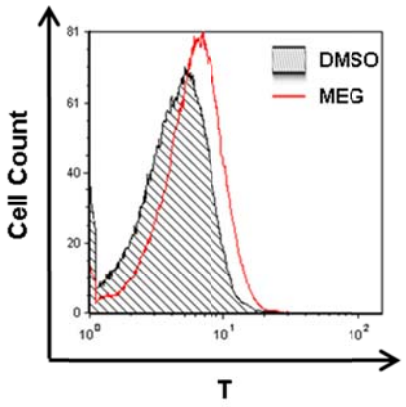


Figure 3.2. Mesendogen (MEG) fails to induce a distinct T^+ mesoderm progenitor population when applied to a differentiation culture alone. FACS analysis of T in H9 hESCs treated with DMSO and 5 μM MEG for 4 days.

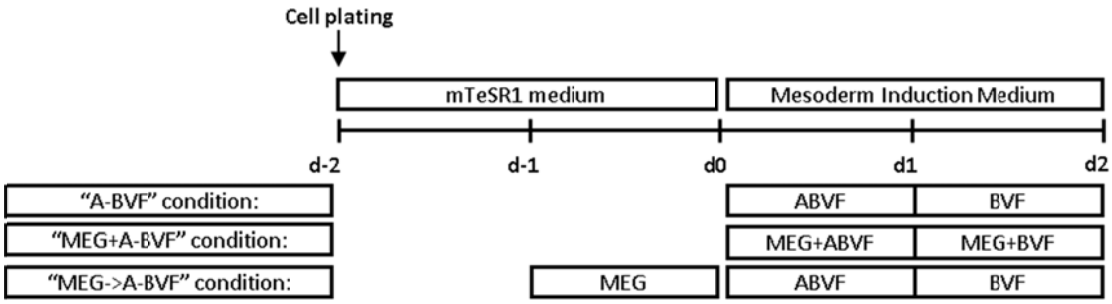


Figure 3.3. Schematic overviews of mesoderm differentiation conditions used in this study.

Seeding Density (x 10^5 cells/cm^2)c	0.1	0.25	0.5	0.75
A-BVF	73.4± 2.3	27.4± 9.0	4.2± 6.9	3.5± 6.0
MEG + A-BVF	82.7± 2.6	77.3± 11.7	61.9± 37.4	45.5± 39.5
MEG -> A-BVF		88.4± 4.0	73.8± 15.3	82.5± 13.2

Figure 3.4. MEG induces nearly homogeneous mesoderm differentiation in combination with the A-BVF condition: FACS summary. Summary of FACS results showing the percentages of T^+EOMES^+ population acquired after 2 days of mesoderm differentiation of H9 and H1 hESCs under various plating densities and differentiation conditions. Results from six independent experiments (n = 6) are shown as mean ± SD; 3 – 4 data points were collected and analyzed for each condition.

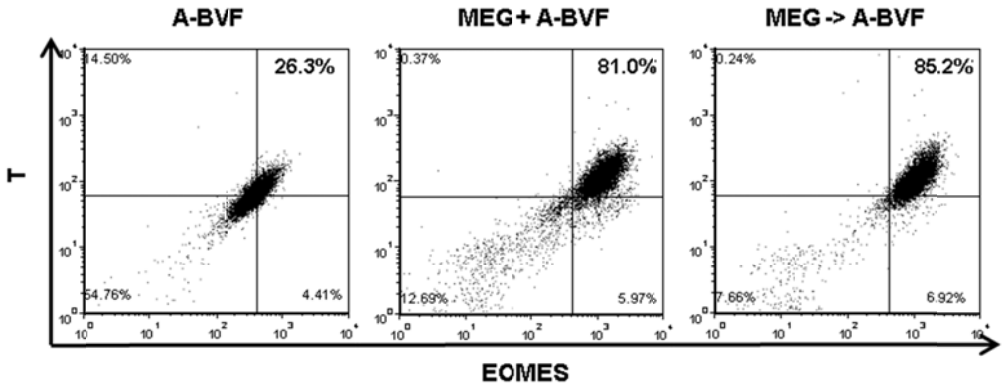


Figure 3.5. MEG induces nearly homogeneous mesoderm differentiation in combination with the A-BVF condition: FACS. FACS analysis of T and EOMES in H9 hESCs treated with mesoderm differentiation conditions including A-BVF, MEG+A-BVF, and MEG->A-BVF. Cells were plated at 0.25×10^5 cells/cm^2.

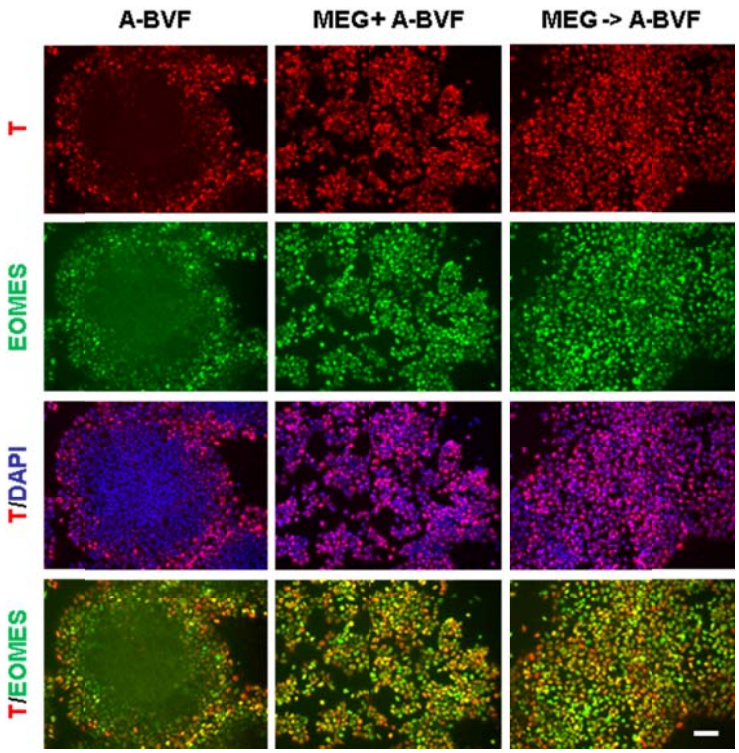


Figure 3.6. MEG induces nearly homogeneous mesoderm differentiation in combination with the A-BVF condition: immunofluorescent staining.
Immunofluorescent staining of T (red) and EOMES (green) in H1 hESCs induced to undergo mesoderm differentiation for ~1.5 days under various conditions. Nuclei were stained with DAPI (blue). Cells were plated at 0.25×10^5 cells/cm^2 for A-BVF condition, 0.75×10^5 cells/cm^2 for MEG+A-BVF condition, and 1.0×10^5 cells/cm^2 for MEG->A-BVF condition.

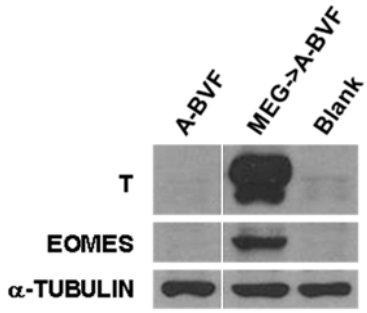


Figure 3.7. MEG induces nearly homogeneous mesoderm differentiation in combination with the A-BVF condition: Western blotting. Western blotting of T and EOMES in H9 hESCs induced to undergo mesoderm differentiation for 2 days under A-BVF and MEG->A-BVF conditions. Cells cultured in mesoderm basal medium without growth factors were used as a control (Blank). Cells were plated at 1.0×10^5 cells/cm^2. α-TUBULIN was used as a loading control.

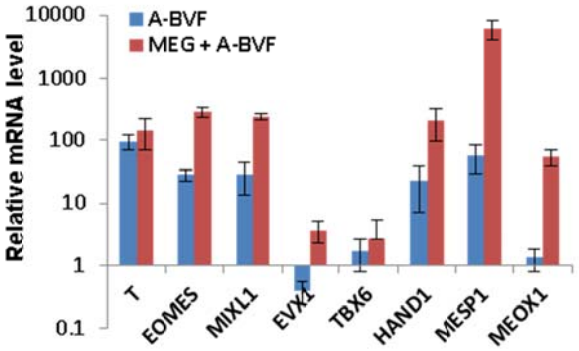


Figure 3.8. MEG induces nearly homogeneous mesoderm differentiation in combination with the A-BVF condition: qPCR. Quantitative-PCR analysis of mesoderm markers in H9 hESCs under A-BVF and MEG+A-BVF conditions for 2 days. All values were normalized to the mRNA level (=1) in cells cultured in mesoderm basal medium without growth factors. Each bar represents the mean ± SD (error bars). n = 2 independent experiments. *ACTB* (*β-actin*) was used as a loading control.

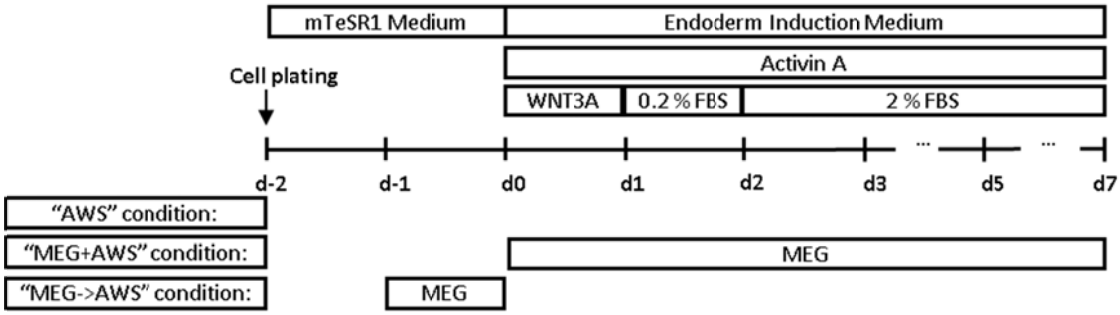


Figure 3.9. Schematic overviews of definitive endoderm differentiation conditions used in this study.

Seeding Density (x 10^5 cells/cm^2)	0.25	0.5	1.0	1.5	2.0
AWS	36.1± 26.2	11.3± 5.6	11.8± 8.7	12.0± 8.3	13.0± 16.8
MEG + AWS		64.9± 24.6	72.7± 19.7	62.6± 25.3	39.1± 28.0
MEG -> AWS			85.9± 6.1	51.7± 37.1	49.1± 37.0

Figure 3.10. MEG induces nearly homogeneous definitive endoderm differentiation in combination with the AWS condition: FACS summary. Summary of FACS results showing the percentages of $SOX17^+FOXA2^+$ population acquired after 7 days of definitive endoderm differentiation of H1 and H9 hESCs under various plating densities and differentiation conditions. Results from six independent experiments (n = 6) are shown as mean ± SD; 3 – 5 data points were collected and analyzed for each condition.

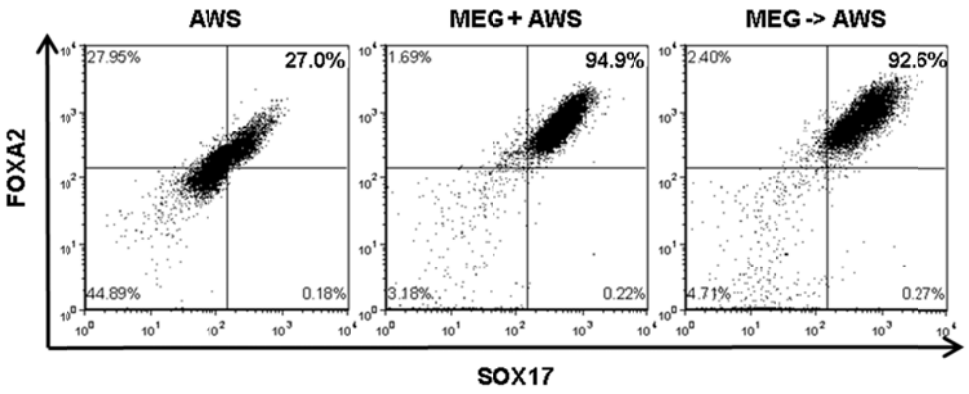


Figure 3.11. MEG induces nearly homogeneous definitive endoderm differentiation in combination with the AWS condition: FACS. FACS analysis of SOX17 and FOXA2 in H1 hESCs treated with definitive endoderm differentiation conditions including AWS, MEG+AWS, and MEG->AWS for 7 days. Cells were plated at 1.0×10^5 cells/cm^2.

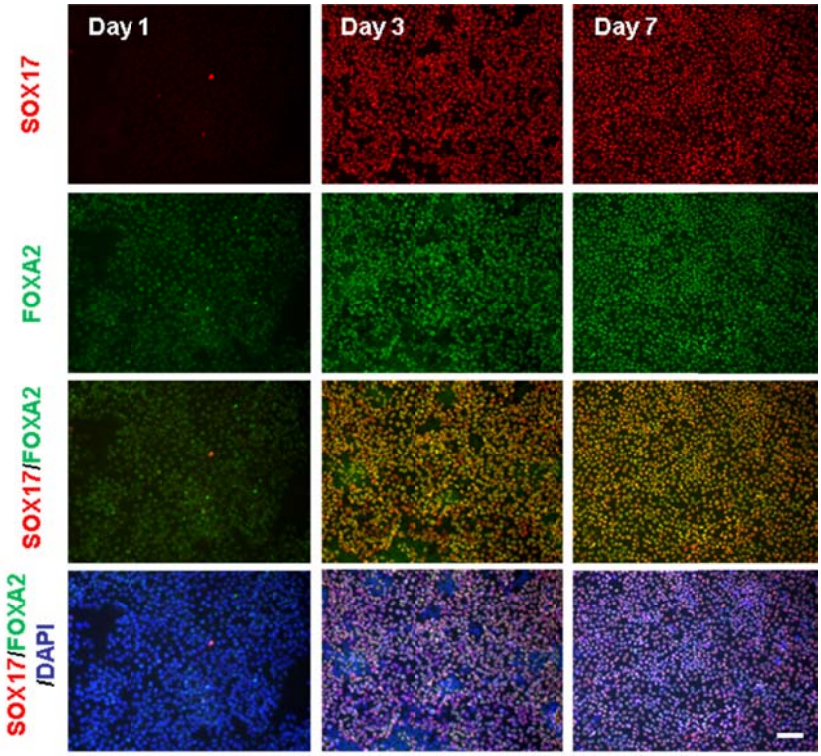


Figure 3.12. MEG induces nearly homogeneous definitive endoderm differentiation in combination with the AWS condition: immunofluorescent staining.
Immunofluorescent staining of SOX17 (red) and FOXA2 (green) in H9 hESCs induced to undergo definitive endoderm differentiation under the MEG->AWS condition for 1, 3, and 7 days. Nuclei were stained with DAPI (blue). Scale bar: 100 μm.

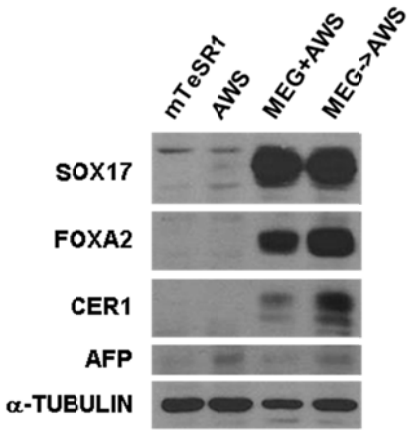


Figure 3.13. MEG induces nearly homogeneous definitive endoderm differentiation in combination with the AWS condition: Western blotting. Western blotting of SOX17, FOXA2, CER1, and AFP in H9 hESCs induced to undergo definitive endoderm differentiation under AWS, MEG+AWS, and MEG->AWS conditions for 7 days. Cells cultured in mTeSR1 medium were used as a control (mTeSR1). Cells were plated at 0.25 $\times 10^5$ cells/cm^2 for the AWS condition, 1.0 $\times 10^5$ cells/cm^2 for the MEG+AWS condition, and 1.5 $\times 10^5$ cells/cm^2 for the MEG->AWS condition. α-TUBULIN was used as a loading control.

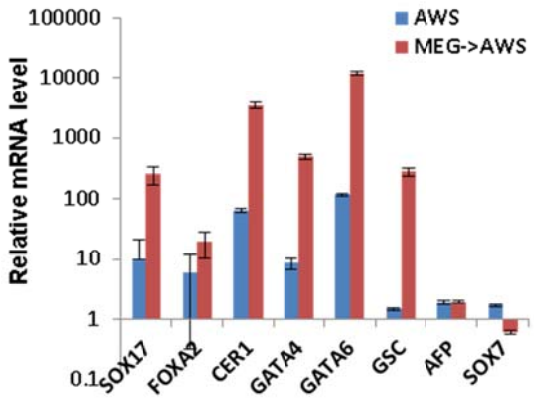


Figure 3.14. MEG induces nearly homogeneous definitive endoderm differentiation in combination with the AWS condition: qPCR. Quantitative-PCR analysis of endoderm markers in H9 hESCs under AWS and MEG->AWS conditions for 5 days. All values were normalized to the mRNA level (=1) in cells cultured in mTeSR1 medium. Each bar represents the mean ± SD (error bars). n = 2 independent experiments. *ACTB* (*β-actin*) was used as a loading control.

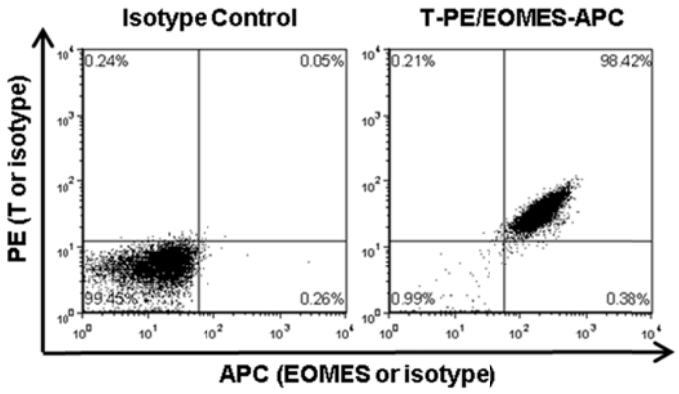


Figure 3.15. MEG induces nearly homogeneous mesoderm differentiation in combination with the A-BVF condition: isotype control. FACS analysis of PE and APC in H9 hESCs cultured in mTeSR1 medium. Cells were stained with T-PE and EOMES-APC antibodies (T-PE/EOMES-APC) or non-specific IgG isotype control antibodies conjugated with the same fluorophores (isotype control). Isotype IgG, when used as a gating control, gave rise to a high percentage of false-positive cells in the undifferentiated mTeSR1 culture (98.42% $PE^{+}APC^{+}$ in the T-PE/EOMES-APC stained sample).

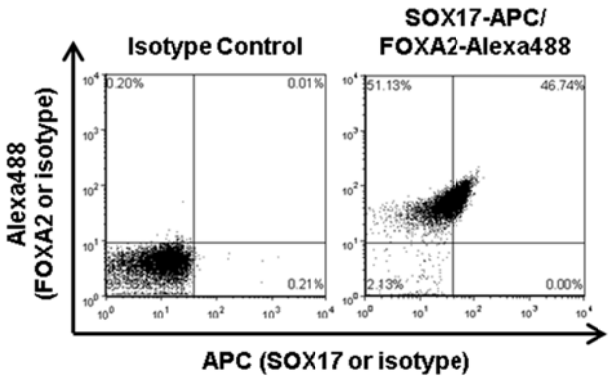


Figure 3.16. MEG induces nearly homogeneous definitive endoderm differentiation in combination with the AWS condition: isotype control. FACS analysis of SOX17 and FOXA2 in H9 hESCs cultured in mTeSR1 medium. Cells were stained with SOX17-APC and FOXA2-Alexa488 antibodies (SOX17-APC/FOXA2-Alexa488) or non-specific IgG isotype control antibodies conjugated with the same fluorophores (isotype control). Isotype IgG, when used as a gating control, gave rise to a high percentage of false-positive cells in the undifferentiated mTeSR1 culture (~50% APC^+/$Alexa488^+$ and ~98% $Alexa488^+$ in the SOX17-APC/FOXA2-Alexa-488 stained sample).

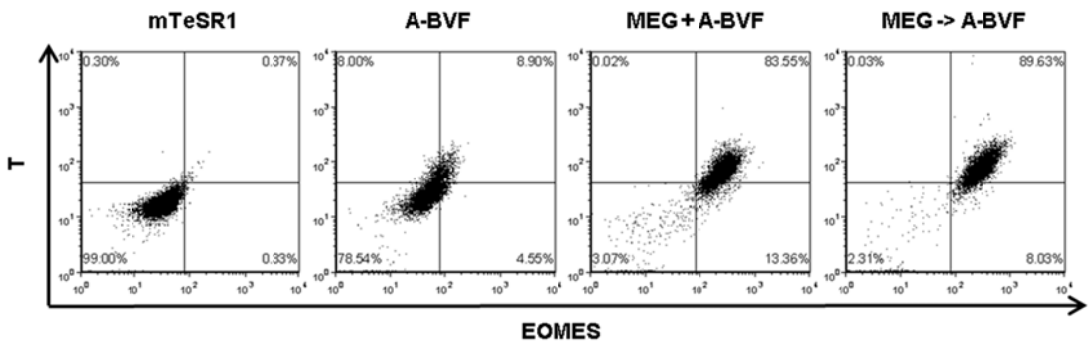


Figure 3.17. MEG induces nearly homogeneous mesoderm differentiation in combination with the A-BVF condition: hiPSC. FACS analysis of T and EOMES in MMW2 hiPSCs differentiated in A-BVF, MEG+A-BVF, and MEG->A-BVF conditions for 2 days. Cells were plated at 0.5×10^5 cells/cm^2 for A-BVF and MEG->A-BVF conditions and 0.25×10^5 cells/cm^2 for MEG+A-BVF condition. Cells cultured in mTeSR1 medium were used as a gating control.

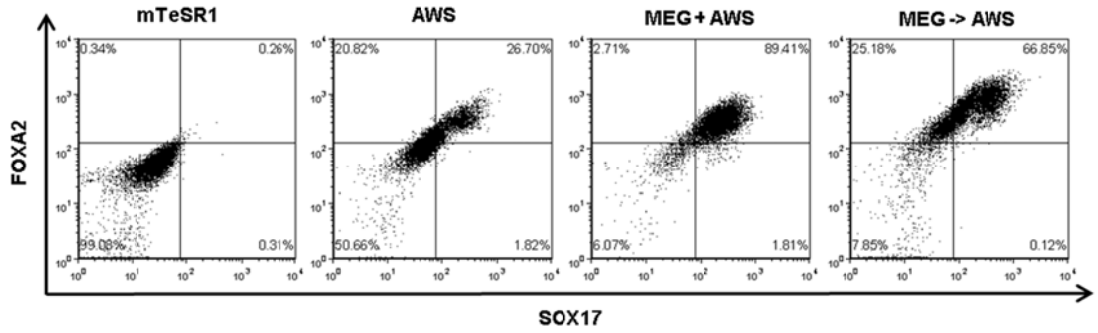


Figure 3.18. MEG induces nearly homogeneous definitive endoderm differentiation in combination with the AWS condition: hiPSC. FACS analysis of SOX17 and FOXA2 in MMW2 hiPSCs cells induced to undergo definitive endoderm differentiation under AWS, MEG+AWS, and MEG->AWS conditions for 7 days. Cells were plated at 1.0×10^5 cells/cm^2 for AWS and MEG+AWS conditions, and at 0.5×10^5 cells/cm^2 for MEG->AWS condition. Cells cultured in mTeSR1 medium were used as a gating control.

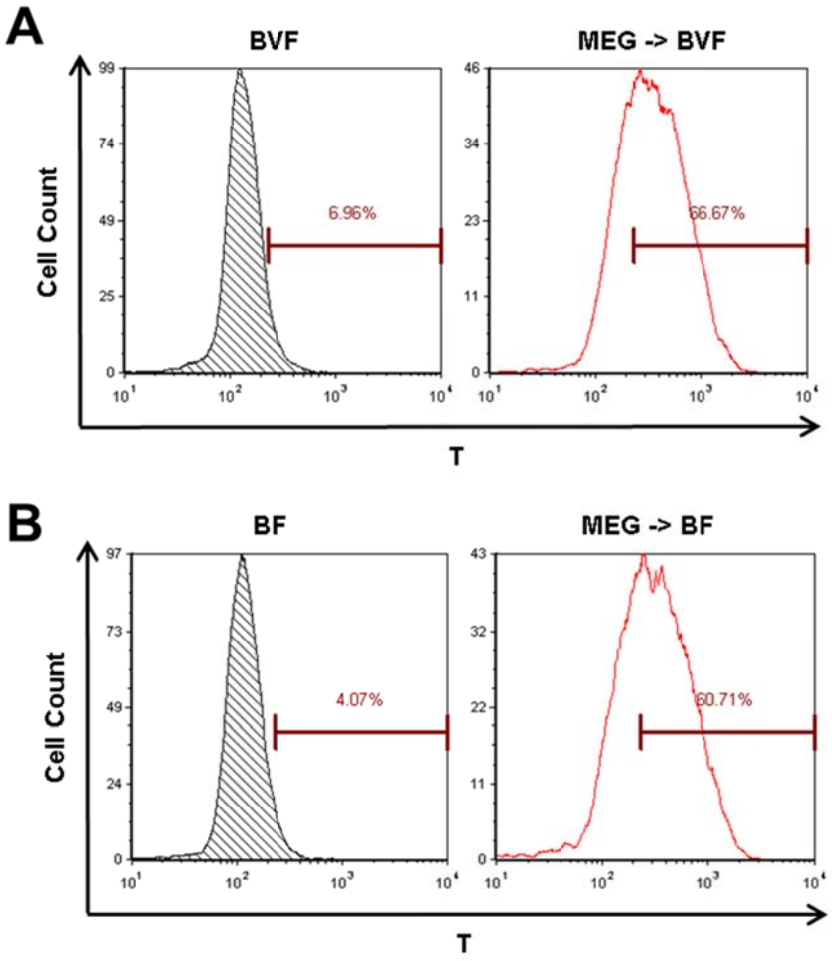


Figure 3.19. MEG induces nearly homogeneous mesoderm differentiation in combination with the A-BVF condition: BVF and BF conditions. (**A**) FACS analysis of T in H1 hESCs differentiated in BMP4 (10 ng/ml), VEGF (10 ng/ml), and bFGF (20 ng/ml) (BVF condition) for 1.5 days, with or without MEG (10 μM) pre-treatment. (**B**) FACS analysis of T in H1 hESCs differentiated in BMP4 (10 ng/ml) and bFGF (20 ng/ml) (BF condition) for 1.5 days, with or without MEG pre-treatment.

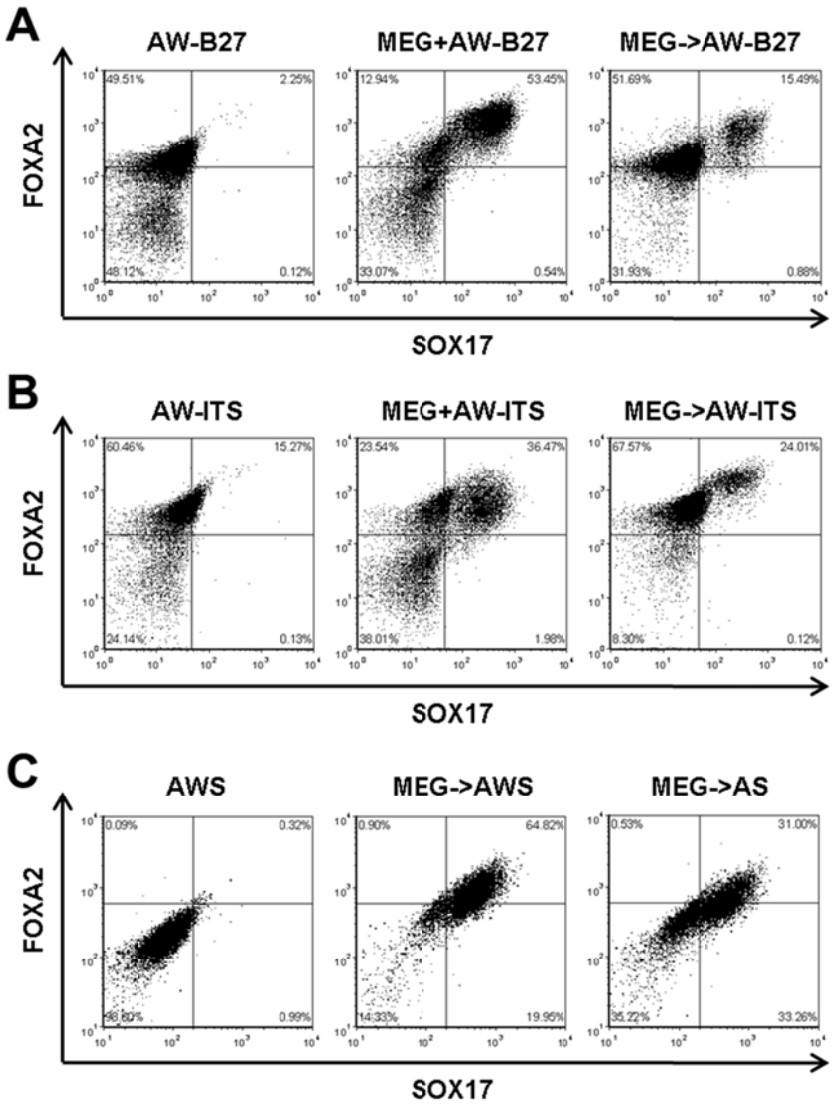


Figure 3.20. MEG induces nearly homogeneous definitive endoderm differentiation in combination with the AWS condition: variations of the AWS condition. (**A**) FACS analysis of SOX17 and FOXA2 in H9 hESCs induced to undergo definitive endoderm differentiation for 5 days in AW-B27, MEG+AW-B27, and MEG->AW-B27 conditions. B27-supplement was used in these conditions instead of serum. Cells were plated at 1.0×10^5 cells/cm^2. (**B**) FACS analysis of SOX17 and FOXA2 in H9 hESCs induced to undergo definitive endoderm differentiation for 5 days in AW-ITS, MEG+AW-ITS, and MEG->AW-ITS conditions. ITS- (insulin, transferrin, and selenium) supplement was used in these conditions instead of serum. Cells were plated at 2.0×10^5 cells/cm^2. (**C**) FACS analysis of SOX17 and FOXA2 in H1 hESCs induced to undergo definitive endoderm differentiation for 7 days in AWS condition, MEG->AWS condition, and MEG->AS condition which lacked WNT3A. Cells were plated at 2.0×10^5 cells/cm^2.

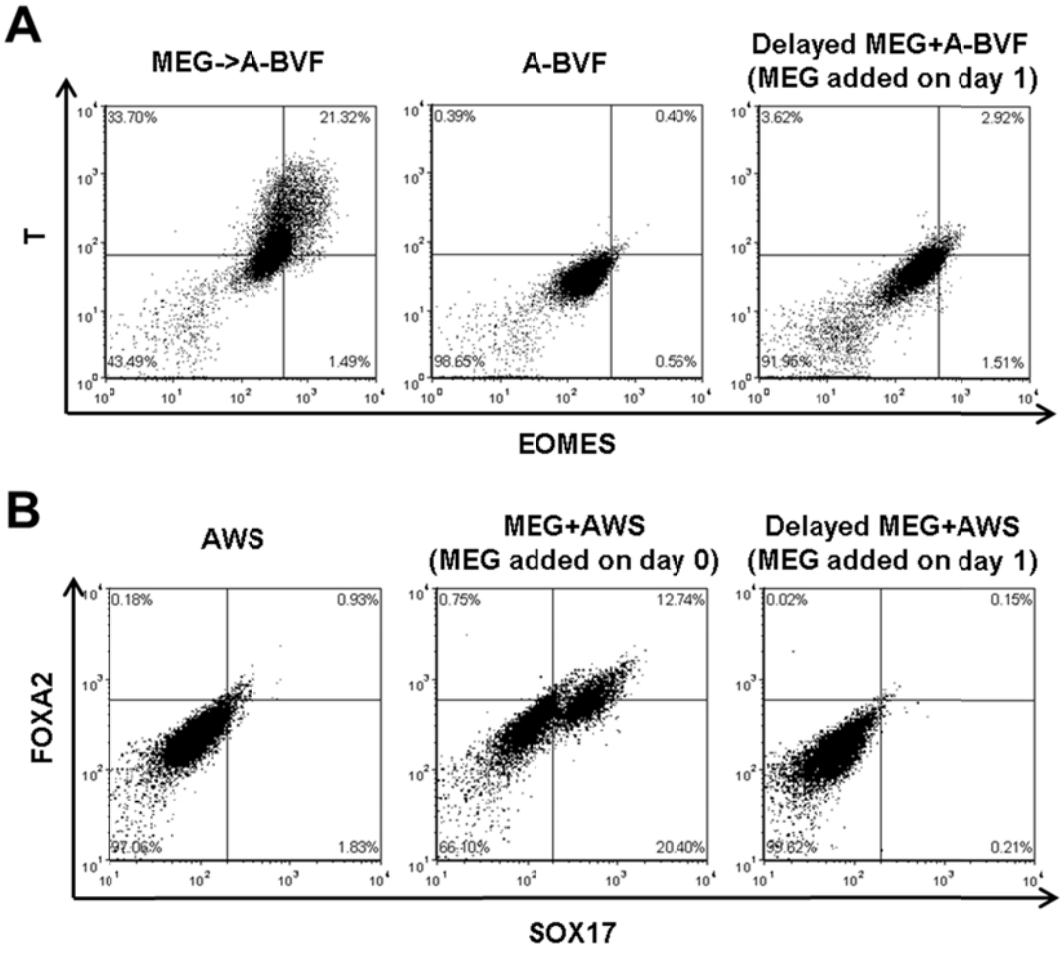


Figure 3.21. Delayed MEG treatment conditions are ineffective. (**A**) FACS analysis of T and EOMES in H9 hESCs differentiated in MEG->A-BVF, A-BVF, and delayed MEG+A-BVF (with MEG added on day 1 instead of day 0) conditions. (**B**) FACS analysis of SOX17 and FOXA2 in H1 hESCs differentiated for 7 days in AWS condition and MEG+AWS condition with MEG added on day 0 or day 1. Cells were plated at 1.5×10^5 cells/cm^2.

A

Percent of Control (%Ctrl)	90 - 100	80 - 90	63 - 80	≤ 50
Number of kinases within each range of %Ctrl	365	64	25	2

B

Hit kinase	%Ctrl
TRPM6	50
NEK3	47

Figure 3.22. A kinome screen identifies potential targets of MEG. (**A**) Summary of KINOMEscan result. Lower values of %Ctrl indicate stronger interactions between the test compound and the kinases. (**B**) %Ctrl values for TRPM6 and NEK3 as detected in KINOMEscan.

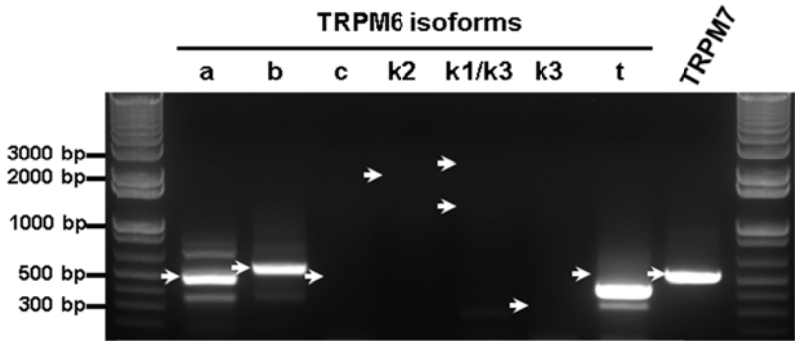


Figure 3.23. TRPM6 isoforms are selectively expressed in undifferentiated hESCs. Agarose gel electrophoresis showing RT-PCR amplification of undifferentiated H1 hESCs using TRPM6 isoform-specific PCR primers. White arrowheads point to positions where a band is supposed to appear if the isoform under detection is present. Note that the band shown in the "t" lane is not in the right position and thus does not indicate expression of the t isoform.

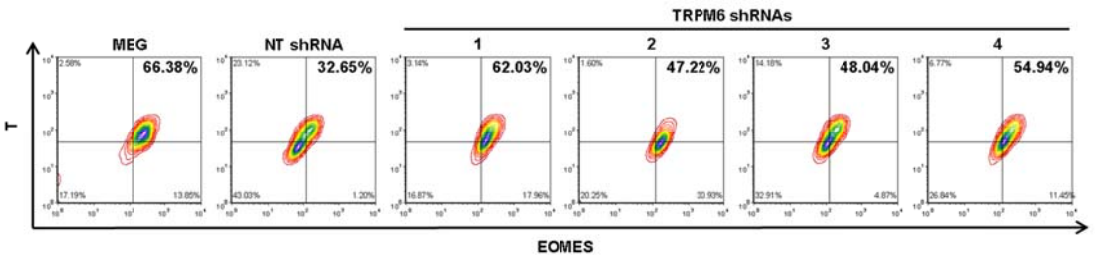


Figure 3.24. TRPM6 knock-down enhances A-BVF-induced mesoderm differentiation. FACS analysis of T and EOMES in H9 hESCs treated with NT shRNA, TRPM6 shRNAs (1 - 4), and MEG (pre-treatment, 10 μM) under mesoderm differentiation condition (A-BVF condition) for 2 days.

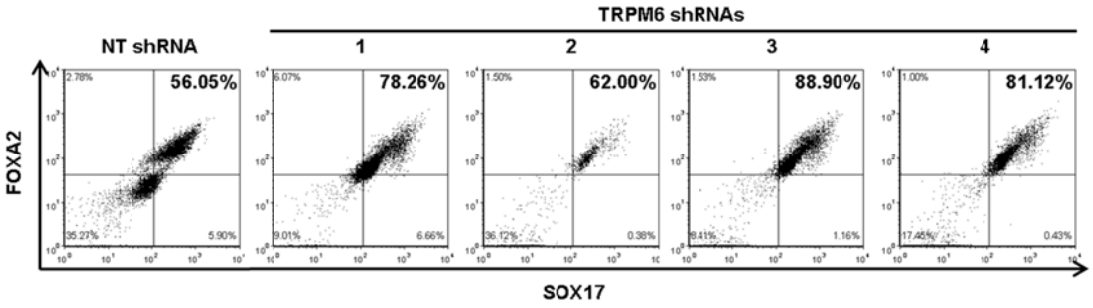


Figure 3.25. TRPM6 knock-down enhances AWS-induced DE differentiation. FACS analysis of SOX17 and FOXA2 in H1 hESCs treated with NT shRNA and TRPM6 shRNAs (1 - 4) under DE differentiation condition (AWS condition) for 7 days.

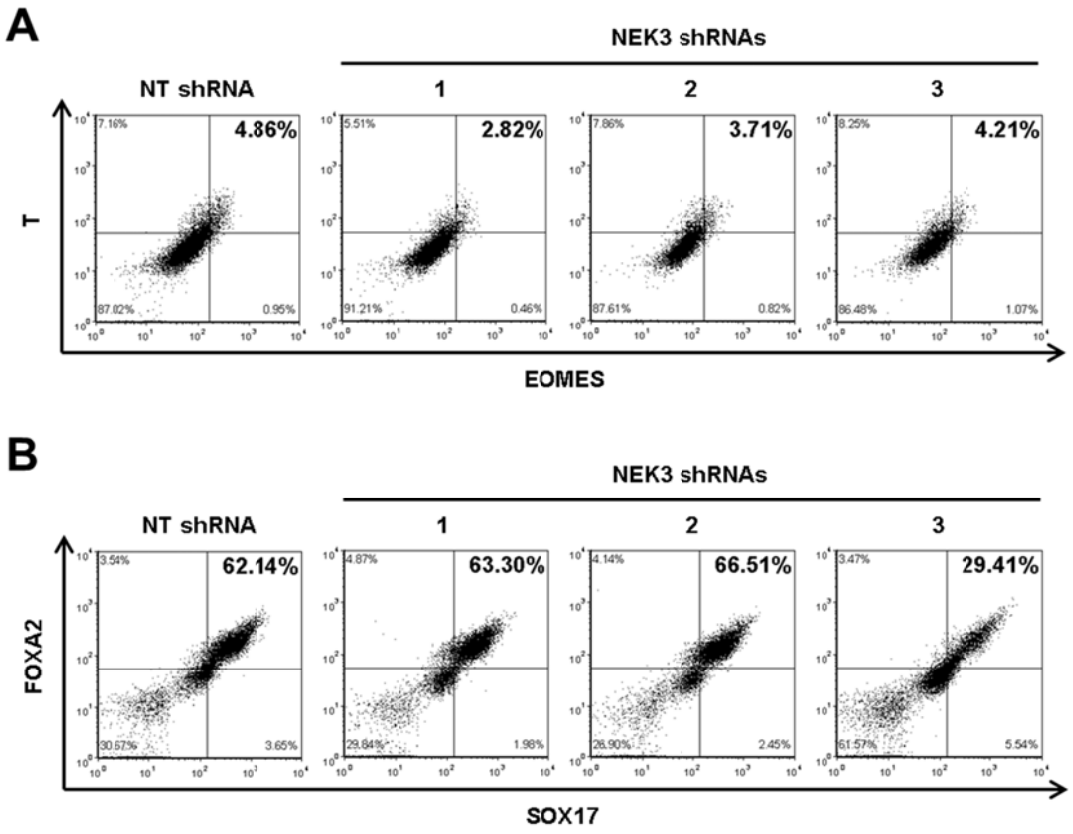


Figure 3.26. NEK3 knock-down shows no effect on differentiation. (**A**) FACS analysis of T and EOMES in H1 hESCs transfected with NT shRNA and NEK3 shRNAs and differentiated under A-BVF condition for 2 days. (**B**) FACS analysis of SOX17 and FOXA2 in H1 hESCs transfected with NT shRNA and NEK3 shRNAs and differentiated under AWS condition for 7 days.

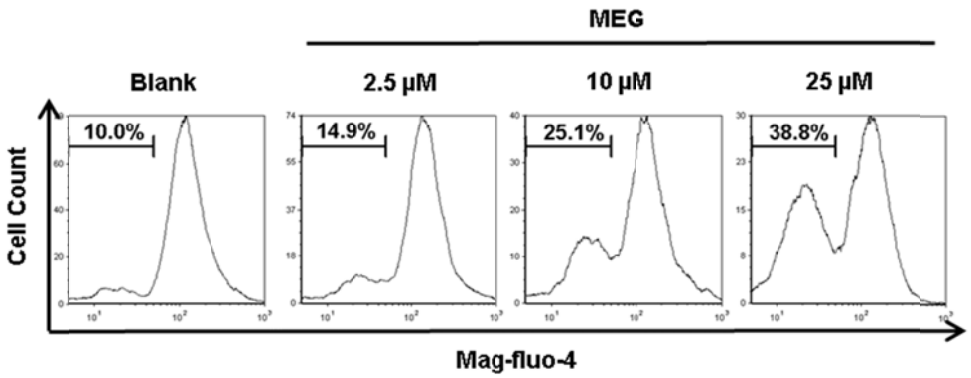


Figure 3.27. MEG reduces intracellular magnesium level. FACS analysis of magnesium sensor Mag-fluo-4 in H1 hESCs untreated or treated with MEG overnight.

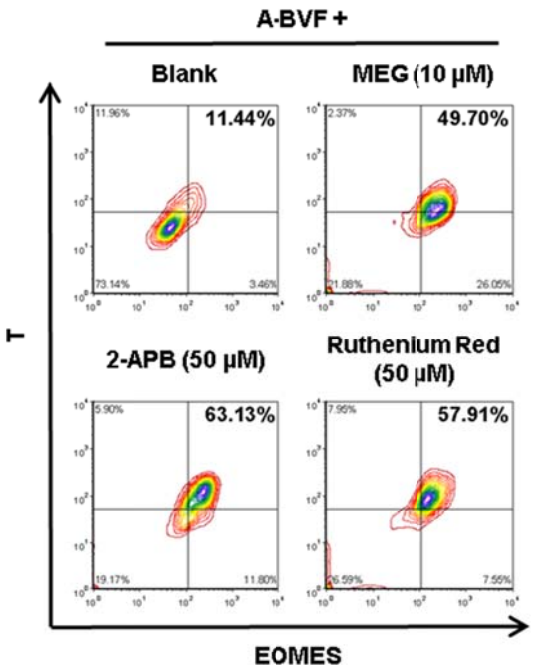


Figure 3.28. TRPM6/TRPM7 channel blockers phenocopy MEG by enhancing A-BVF differentiation efficiency. FACS analysis of T and EOMES in H1 hESCs untreated or pre-treated with MEG (10 μM), Ruthenium Red (50 μM), and 2-APB (50 μM) under A-BVF condition for 2 days.

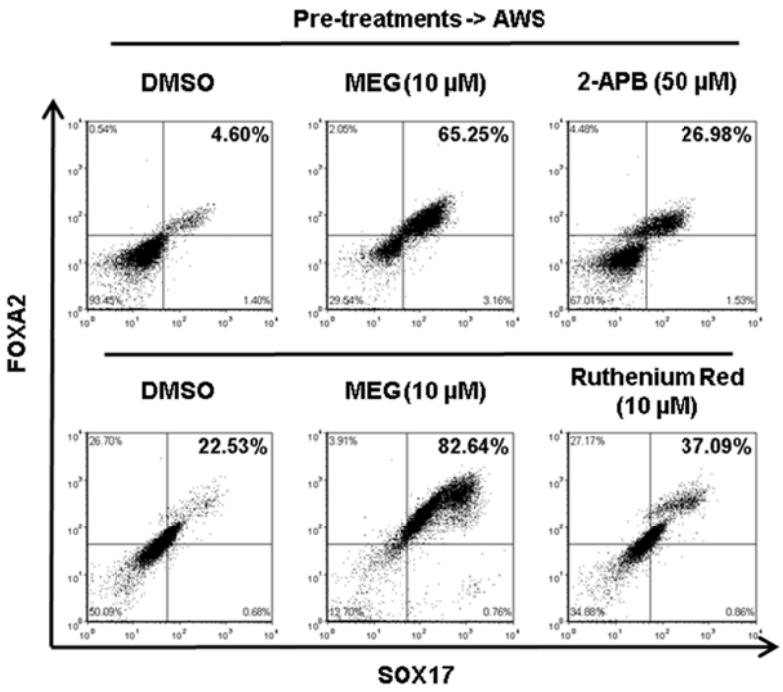


Figure 3.29. TRPM6/TRPM7 channel blockers phenocopy MEG by enhancing AWS differentiation efficiency. FACS analysis of SOX17 and FOXA2 in two independent experiments of H1 hESCs pre-treated with DMSO, 10 μM MEG, and 50 μM 2-APB (top panel) and DMSO, MEG, and 10 μM Ruthenium Red (bottom panel) under definitive endoderm differentiation condition (AWS condition) for 7 days.

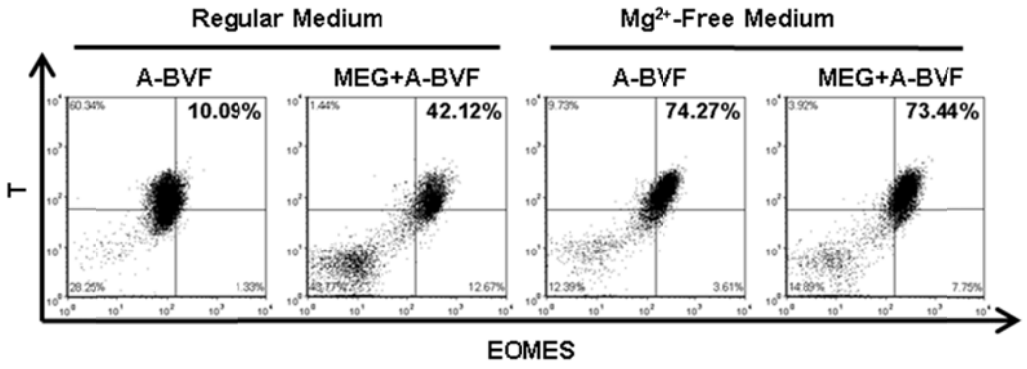


Figure 3.30. Manipulations of Mg^{2+} level in culture medium affect the efficiency of A-BVF differentiation. FACS analysis of T and EOMES in H1 hESCs untreated or treated with MEG (10 μM) under A-BVF differentiation condition in regular medium or Mg^{2+}-free medium for 2 days.

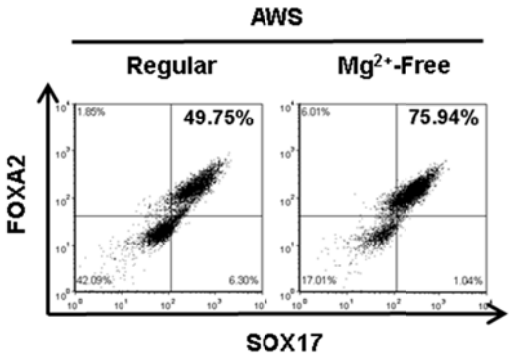


Figure 3.31. Mg^{2+}-withdrawal phenocopies MEG by enhancing AWS differentiation efficiency. FACS analysis of SOX17 and FOXA2 in H1 hESCs undergoing DE differentiation in AWS condition, using regular medium or Mg^{2+}-free medium, for 7 days.

Tables

Antibody	Source	Catalog Number	Dilution
OCT4	Santa Cruz	sc-9081	1:1,000 (WB)
SOX2	Millipore	AB5603	1:500 (WB)
NANOG	Cell Signaling	4903P	1:500 (WB)
α-TUBULIN	Abcam	ab11304	1:10,000 (WB)
GAPDH	GeneScript	A00192	1:10,000 (WB)
T	R&D systems	AF2085	1:500 (WB/IF)
EOMES	Abcam	ab23345	1:500 (WB/IF)
SOX17	R&D systems	AF1924	1:500 (WB/IF)
FOXA2	Millipore	07-633	1:500 (WB/IF)
CER1	Thermo-Pierce	MA5-15554	1:500 (WB)
AFP	DAKO	A0008	1:500 (WB)
SOX1	R&D systems	AF3369	1:500 (WB)
PAX6	Covance	PRB-278P	1:500 (WB)
T-PE	R&D Systems	IC2085P	1:20 (FACS)
EOMES-APC	R&D Systems	IC6166A	1:20 (FACS)
FOXA2-Alexa488	R&D Systems	IC2400G	1:40 (FACS)
SOX17-APC	R&D Systems	IC1924A	1:20 (FACS)
Goat IgG-PE isotype control	R&D Systems	IC108P	1:20 (FACS)
Goat IgG-APC isotype control	R&D Systems	IC108A	1:20 (FACS)
Mouse IgG_{2B}-APC isotype control	BioLegend	400319	1:20 (FACS)
Goat IgG-Alexa488 isotype control	R&D Systems	IC108G	1:40 (FACS)

(WB: Western blotting; IF: immunofluorescence)

Table 3.1. Antibody sources and dilutions

Primer Name	Forward Primer (5'-3')	Reverse Primer (5'-3')
ACTB	agagctacgagctgcctgac	cgtggatgccacaggact
OCT4-Isoform-A	cttctcgcccccctccaggt	aaatagaacccccagggtgagc
NANOG	tttggaagctgctggggaag	gatgggaggaggggagagga
SOX2	ggcagctacagcatgatgcaggagc	ctggtcatggagttgtactgcagg
T (Brachyury)	gctgtgacaggtacccaacc	catgcaggtgagttgtcagaa
EOMES	gtggggaggtcgaggttc	tgttctggaggtccatggtag
MIXL1	ggcgtcagagtgggaaatcc	gcagttcacatctacctcaagag
EVX1	ttcacccgagagcagattg	ccggttctggaaccacac
TBX6	gaacggcagaaactgtaagagg	gtgtgtctccgctcccatag
HAND1	aaaggccctacttccagagc	tgcgctgttaatgctctcag
MEOX1	aaagtgtcccctgcattctg	cactccagggttccacatct
MESP1	ctgttggagacctggatgc	cgtcagttgtcccttgtcac
SOX17	acgccgagttgagcaaga	tctgcctcctccacgaag
FOXA2	tgggagcggtgaagatggaagggcac	tcatgccagcgcccacgtacgacgac
CER1	acagtgcccttcagccagact	acaactacttttttcacagccttcgt
GSC	gaggagaaagtggaggtctggtt	ctctgatgaggaccgcttctg
GATA4	ggaagcccaagaacctgaat	gttgctggagttgctggaa
GATA6	aatacttcccccacaacacaa	ctctcccgcaccagtcat
AFP	agcttggtggtggatgaaac	ccctcttcagcaaagcagac
SOX7	ctcagggcagggaggtct	gcactcggataaggagagtcc

Table 3.2. Names and sequences of qPCR primers

Target	shRNA	Sequence (5’ – 3’)
TRPM6	TRPM6-1	ccggcctggcataaagaatgtatatctcgagatatacattctttatgccaggttttt
	TRPM6-2	ccggccctctaatctaaagcgagttctcgagaactcgctttagattagagggttttt
	TRPM6-3	ccggccaaattctaatggagtgtatctcgagatacactccattagaatttggttttt
	TRPM6-4	ccgggctccctatctgataactcaactcgagttgagttatcagatagggagcttttt
NEK3	NEK3-1	ccgggcagtcccatagaacagaaatctcgagatttctgttctatgggactgctttt
	NEK3-2	ccggccttattatgtgcctccagaactcgagttctggaggcacataataaggttttt
	NEK3-3	ccggcgaagcataacacaccaagaactcgagttcttggtgtgttatgcttcgttttt

Table 3.3. Names and sequences of shRNAs

CHAPTER 4: SPONTANEOUS SELF-ORGANIZATION OF A PRIMITIVE LYMPHATIC PLEXUS-LIKE STRUCTURE FROM HUMAN EMBRYONIC STEM CELLS TRIGGERED BY SMALL MOLECULES

Abstract

Human embryonic stem cells (hESCs), especially the emerging models of hESC-derived self-organizing organoids, provide a potential platform for studying early human development and modeling human developmental disorders. However, previously reported organoid systems rely on pre-designed programs of external guidance to direct their differentiations, and were thus unable to reveal truly intrinsic developmental programs embedded in hESCs. Here I report two novel small molecules (Lymphgen 1/2) that trigger an unguided and spontaneous self-organizing event which gives rise to a balloon-like organoid structure. Gene expression analyses, functional assays, and morphological studies demonstrated that this self-organizing event recapitulates the *in vivo* human lymphatic morphogenesis program. Using this system, I unveiled the unguided emergence of a DiI-Ac-LDL^{+}VE-cadherin^{+}CD31^{+}CD34^{+}KDR^{+}CD43^{-} endothelial progenitor population, and identified a previously unknown VEGFR3^{+} progenitor population which may be responsible for human embryonic lymphatic development. This system provides a rare opportunity to visualize a truly spontaneous human developmental program *in vitro*.

Introduction

The mechanisms of organogenesis remain a central piece of the puzzle for human embryonic development. Traditional differentiation methods such as embryoid body (EB) formation (Itskovitz-Eldor et al., 2000a) and growth factor-directed 2-dimensional (2-D)

differentiations (Borowiak et al., 2009a) were unable to recapitulate 3-dimensional (3-D) self-organizing events during organogenesis. To solve this problem, the Yoshiki Sasai group pioneered the field of 3-D organoid self-organization by developing a series of conditions that guide hESCs into developmental programs mimicking various aspects of early human neurogenesis (Eiraku et al., 2011; Eiraku et al., 2008; Muguruma et al., 2010; Nakano et al., 2012; Sasai, 2013; Suga et al., 2011). Several groups later developed their own approaches for 3-D organoid-formation which recapitulated organogenesis events of organs originated from all three germ layers (Antonica et al., 2012; Cao et al., 2011; Lancaster et al., 2013; Mariani et al., 2012; Saito et al., 2011; Ueda et al., 2010; Wang and Ye, 2009; Xia et al., 2013). These studies all adopted a suspension culture system similar to that used for EB-formation. A different approach to organoid self-formation was later developed for the self-organization of intestine-like (McCracken et al., 2014; Spence et al., 2011; Wells and Spence, 2014) or kidney-like (Takasato et al., 2014) 3-D organoids that grew out of 2-D cultures. This "2-D" approach presents a unique advantage over the "3-D" approach in that it allows researchers to easily visualize and track the process of "sprouting", the formation of a 3-D luminal structure out of a 2-D epithelial sheet (Pohl et al., 2000), which is a vital process for a variety of organogenesis events such as vasculogenesis (Gerhardt, 2008), tracheal branching (Ghabrial and Krasnow, 2006), ureteric-bud branching (Basson et al., 2006), and breast development (Sternlicht, 2006). However, a major limitation of all reported self-organizing systems is that they all require continues guidance of external differentiation factors, such as growth factors, throughout their courses of differentiation, and are thus guided rather than truly spontaneous self-organizing processes.

A delicate state of balance exists between the pluripotency regulatory network and the differentiation regulatory pathways, and is required for the maintenance of pluripotency (Boyer et al., 2005; Shu et al., 2013). Disrupting this balanced state by altering the expressions or activities of different components that are important for the maintenance of this balance will trigger hESC differentiations in different ways. For example, inhibition of HSPA8 by a small molecule Displg causes un-directional differentiation (Chapter 2), whereas another small molecule MEG was shown to specifically enhance mesoderm and definitive endoderm differentiations (Chapter 3). I thus hypothesized that, by applying a stimulus that undermines this delicate balance in a certain way during the initial phase of differentiation, I may be able to trigger a truly spontaneous self-organizing event that can propagate on its own by following an intrinsic organogenesis program embedded in hESCs, without the need for any external guidance.

The key to the success of this proposal is to find a proper stimulus that is able to trigger such a delicate differentiation event. Due to the unpredictable nature of this hypothesized mechanism of induction, a large number of highly diversified stimuli need to be tested in order for me to identify such a stimulus: no other method achieves this goal better than small molecule screening. Small molecules provide a source of practically inexhaustible structural diversity - a multitude of bioactive small molecules have been reported to modulate a large variety of biological molecules in various biological pathways - and are thus ideal candidate stimuli for my study.

Here I report the identification of two novel small molecules that trigger an unguided and spontaneous self-organization program during hESC differentiation, giving rise to a balloon-like 3-D organoid structure. This balloon-like structure was then identified as an *in vitro* counterpart of primitive lymphatic plexus which was the building blocks of embryonic lymphatic development *in vivo*, and the balloon-formation process was thus found to recapitulate lymphatic morphogenesis during human embryonic development. These two compounds were therefore named Lymphgen 1 and 2, respectively, for their properties as lymphatic-inducing reagents. Using this system I identified a spontaneously emerged population of DiI-Ac-LDL^{+}VE-cadherin^{+}CD31^{+}CD34^{+}KDR^{+}CD43^{-} endothelial progenitor cells, discovered a previously uncharacterized VEGFR3^{+} progenitor population which may represent lymphatic progenitors responsible for human embryonic lymphatic development, and visualized the truly spontaneous progression of human lymphatic morphogenesis *in vitro*.

Results

Morphological screening identified two structurally similar compounds that induce a self-organizing balloon-like structure

I previously conducted a high throughput screening (HTS) study and found 29 bioactive small molecules that potently disrupt hESC pluripotency (Chapter 2). Using this compound collection I conducted a small-scale morphological screen (Figure 4.1 and Materials and Methods) to search for small molecules that can induce self-organization of 3-D organoid structures. Briefly, hESCs were plated as 2-D cultures so that the emergence of any 3-D structures could be readily visible. Two consecutive phases were included in the screen: for the “induction phase”, small molecules were applied to hESCs

as the initial stimuli for periods long enough to induce differentiations; then for the "spontaneous self-organization phase", the molecules were withdrawn, and the cultures allowed to go through self-driven differentiation processes in the hope that 3-D self-organizing organoids may emerge from some of the cultures. Both phases were growth factor-free.

Using this screen I identified two structurally similar compounds 5373471 and 6018575 (Figure 4.2) as inducers of a self-organizing balloon-like structure (Figure 4.3A; hereafter referred to as "balloons"). For balloon formation, 5373471 (2.5 - 5 μM) or 6018575 (2.5 - 5 μM) was first added for 5 - 6 days to induce differentiation; cells were then let grow freely without any further guidance (Materials and Methods). Balloons generally start to emerge at around day 10. A typical "balloon" consists of a closed spherical structure which is usually suspended in the culture medium (but sometimes remained adherent to the culture surface) and a stalk-like structure ("stalks" hereafter) which links the spherical structure to the culture surface (Figures 4.3). The spherical structure in a balloon is always filled with colorless fluid, and it keeps accumulating fluid and bulge during prolonged periods of culture. The suspended balloons gently move about, as driven by disturbances in the culture medium, while remaining firmly attached to the culture surface. I hereafter refer to compounds 5373471 and 6018575 as Lymphgen 1 and Lymphge 2, respectively, for simplicity and consistency throughout the text.

Gene expression profiling and functional tests revealed an endothelial identity of the balloons

To identify lineage-specificity of the balloons, I examined their gene-expression profile. Quantitative-PCR (qPCR) analysis showed consistent elevations of endothelial lineage-specific markers, including general endothelial markers VE-cadherin (Vestweber, 2008), KDR (Samuel et al., 2013), CD31 (Kane et al., 2010), CD34 (Nakajima-Takagi et al., 2013), etc., and markers specific to arterial, venous, and lymphatic endothelia, in the balloon cultures as compared to the DMSO controls (Figures 4.4). Fluorescent-activated cell sorting (FACS) also showed that endothelial-specific markers VE-cadherin, CD31, CD34, CXCR4 (Gupta et al., 1998), UEA1 (a human endothelial cell-specific lectin) (Christenson and Stouffer, 1996), and KDR were elevated compared to the DMSO control (Figure 4.5). Only the spherical parts of the balloons (Figure 4.6, red arrows) and the core of the stalks (Figure 4.6, white arrows) stained positive for VE-cadherin and actively took up DiI-Ac-LDL (Figure 4.6), both of which characteristics of endothelial cells (Vestweber, 2008; Voyta et al., 1984). To rule out the existence of other lineages in the balloon cultures, I examined two lineages in particular that are known to be closely related to the endothelial lineage during development: the mural cells (including vascular smooth muscle cells and pericytes) (Sone et al., 2007) and the blood lineage (Kennedy et al., 2007). The mRNA expressions of blood lineage-specific markers, including hematopoietic progenitor markers and sub-lineage markers specifying lymphoid and myeloid identities, were not elevated compared to the DMSO control (Figure 4.7); neither was the mural cell-specific marker Calponin (Figure 4.7). Altogether, these results suggested an endothelial identity of the balloons, especially for the core parts of the

balloon structure, and also indicated that differentiations toward the blood lineage and to mural cells were not triggered in the balloon cultures.

Functional Assays further confirmed the endothelial identity of the balloons. A selective-adhesion assay (Materials and Methods) showed that, 3 days after re-plating of 12-day balloon cultures and a DMSO control at the same density into an endothelial-specific culture condition, the number of adherent cells in cultures seeded by Lymphgen-treated cells was significantly higher than in the culture seeded by the DMSO control (Figure 4.8A). Nearly all remaining cells took up DiI-Ac-LDL (Figure 4.8B), verifying the effectiveness of this selective adhesion method. A Matrigel vessel-formation assay (Materials and Methods), which examines the ability of cells to form vascular networks *in vitro*, also showed that only cells from the balloon cultures formed vessel networks (Figure 4.9).

Morphological analysis gave rise to a working model of balloon self-organization

To study the balloon-formation process, I started with the observation that balloon-like structures start to appear at around day 10, and then continue to emerge and grow for prolonged periods of time. This means that at any given time after day 10, a Lymphgen-induced culture should contain multiple balloon-like structures reaching various stages of balloon-formation. I thus examined multiple cultures at fixed time points after day 10 (with a focus on days 12 - 13) to reconstruct the balloon-formation process.

The general appearance of a culture at around days 12 to 13 is highly heterogeneous. Parts of the culture remain flat on the surface, with some areas maintaining an epithelial appearance (similar to that of a pluripotent colony) while other areas a mesenchymal appearance (Figure 4.10, left column). Other parts of the culture are elevated in the form of balloon-like structures or irregularly shaped cell aggregates (Figure 4.10, middle and right columns). To distinguish endothelial structures from the rest of the culture, I incubated live cultures with DiI-Ac-LDL to allow cells of endothelial identity to take up the dye before microscopic examination. I identified the following distinctive structures formed by DiI-Ac-LDL$^+$ cells. First, I found clusters of DiI-Ac-LDL$^+$ round cells with mesenchymal characteristics that are elevated above (but remain attached to) the epithelial-looking cell sheet (Figure 4.10, left column; hereafter referred to as "clusters"). Second, I found closed spherical structures that remain directly (without the help of "stalks") attached to the culture surface and are much smaller in size compared to the suspended balloons (Figure 4.10, middle column; hereafter referred to as "spheres"). Most of those "spheres" seemed to have grown out of irregularly shaped DiI-Ac-LDL$^+$ cell aggregates (Figure 4.10, middle and right columns, white arrows) as judged by their proximity. Finally, as expected, I identified balloon-like structures (Figure 4.10, right column) as already discussed.

Based on these observations I propose a working model for the balloon-development process (Figure 4.11): DiI-Ac-LDL$^+$ cells, which are likely to be endothelial progenitor cells, first emerge from the epithelial-looking cell sheets and form "clusters";

the clusters then grow to form irregularly-shaped cell aggregates, from which sprout the "spheres"; the spheres eventually grow into balloons (Figure 4.11).

Identification of a DiI-Ac-LDL$^+$VE-cadherin$^+$CD31$^+$CD34$^+$KDR$^+$CD43$^-$ endothelial progenitor population during balloon formation

To verify this model in detail, I first investigated the identity of the cells in "clusters". By double-staining a balloon culture using DiI-Ac-LDL and VE-cadherin, I found that a majority of the cells in the clusters were DiI-Ac-LDL$^+$VE-cadherin$^+$ (Figure 4.12). I then further analyzed the identity of the VE-cadherin$^+$ cells using flow cytometry, and found a distinct VE-cadherin$^+$ population which typically constitutes 0.2 – 0.8% of the total population of a 12- to 13-day balloon culture (Figure 4.13A, top-left plot). By gating the VE-cadherin$^+$ population, I found that almost all VE-cadherin$^+$ cells were positive for endothelial progenitor markers KDR (95.96%) and CD31 (98.16%) (Figure 4.13A, top row). When plotted against the side scatter channel (SSC), CD31$^+$ cells appear as a distinct population as well; this population is almost entirely VE-cadherin$^+$ (Figure 4.13A, bottom row, 93.50%), indicating the existence of a VE-cadherin$^+$CD31$^+$ population in the balloon cultures. To further verify the endothelial identity of this population, I triple stained CD31 with CD34, a marker expressed in both endothelial and hematopoietic progenitors (Nakajima-Takagi et al., 2013), and CD43, a marker restricted to the hematopoietic lineage (Vodyanik et al., 2006). I found that the VE-cadherin$^+$CD31$^+$ population was CD34$^+$, but CD43$^-$ (Figure 4.13B), thus further demonstrating an endothelial rather than hematopoietic identity of this population. Taken together with the staining result, these results demonstrate that the "clusters" contain a population of DiI-

Ac-LDL^{+}VE-cadherin^{+}CD31^{+}CD34^{+}KDR^{+}CD43^{-} cells that are likely to be of an endothelial progenitor identity.

The balloon self-organizing system recapitulates lymphatic morphogenesis in vitro

I then compared the balloon-formation process with *in vivo* endothelial development programs including vascular- (blood vessel) and lymphatic-lineage specifications in order to identify the *in vivo* relevance of balloon morphogenesis, and found striking morphological similarities between balloon formation and lymphatic development. First, lymphatic vessels develop from sac-like structures called primitive lymphatic plexuses (Martinez-Corral and Makinen, 2013; Tammela and Alitalo, 2010) which are morphologically very similar to the "spheres" (Figure 4.10, middle column). Second, primitive lymphatic plexuses form by sprouting out of venous vessels, and as one matures, its connection with the venous vessel would constrict, fuse, and eventually sever (Tammela and Alitalo, 2010); this is also very similar to the balloon-formation process during which a balloon, although never completely detaches from the culture surface, separates itself from the surface by developing a stalk-like structure containing a core that stains positive for DiI-Ac-LDL and VE-cadherin (Figure 4.6, white arrows) but is completely fused and does not contain any cavities. This phenomenon, while hard to be explained under a blood-vessel development scenario, fits well with a lymphatic development model. Finally, mature balloons continue to accumulate fluid and bulge, similar to mature lymphatic vessels which collect interstitial fluid and produce lymph. These morphological similarities strongly suggest that balloon formation mimics the *in vivo* lymphatic morphogenesis.

I then further compared the similarities between balloon formation and lymphatic development on a molecular level. First, I studied resemblances between sphere formation and lymphatic sprouting. Based on *in vivo* studies showing that sprouting primitive lymphatic plexuses express the surface receptor VEGFR3 (Tammela and Alitalo, 2010), I originally hypothesized that, due to the morphological resemblance between spheres and primitive lymphatic plexus, the spheres may represent an *in vitro* counterpart of the primitive lymphatic plexus, and thus stain positive for VEGFR3 *in toto*. Surprisingly, in each sphere, only some of the cells stain positive for VEGFR3 while the rest remain $VEGFR3^-$ (Figure 4.14). Moreover, these $VEGFR3^+$ cells are grouped together in each sphere in the shape of a sprout (Figure 4.14, white arrows; hereafter referred to as the "$VEGFR3^+$ sprouts") which appears to be budding out of the rest of the sphere. I thus adjusted my original hypothesis and proposed that it is the $VEGFR3^+$ sprouts, rather than the spheres *in toto*, that mimic primitive lymphatic plexuses. Second, I compared balloons with the primitive lymphatic plexuses after venous separation. While primitive lymphatic plexuses continue to express VEGFR3 after venous separation, the spherical structures of balloons are also positive for VEGFR3 in their entirety (Figure 4.15), once again demonstrating the lymphatic identity of the balloons. This data also suggested that the balloons most likely originated from the $VEGFR3^+$ sprouts rather than from the $VEGFR3^-$ parts of the spheres.

Finally, I investigated the origin of the $VEGFR3^+$ sprouts. I argued that the $VEGFR3^+$ sprouts may have originated from $VEGFR3^+$ progenitor cells that existed as a

subgroup within the DiI-Ac-LDL^{+}VE-cadherin^{+}CD31^{+}CD34^{+}KDR^{+}CD43^{-} endothelial progenitor population. Indeed, using immunofluorescence staining, I found VEGFR3^{+} cells in the DiI-Ac-LDL^{+}VE-cadherin^{+}CD31^{+}CD34^{+}KDR^{+}CD43^{-} progenitor clusters (Figure 4.16). FACS analysis also showed a VEGFR3^{+} population that only exists in the balloon cultures (Figure 4.17A, black arrow). Double staining of VEGFR3 and CD31 followed by FACS analysis demonstrated that approximately 60% of the CD31^{+} cells in a 12-day balloon culture are positive for VEGFR3 (Figure 4.17B). VEGFR3^{+} progenitor cells have been reported to be responsible for mouse embryonic lymphatic differentiation (Kono et al., 2006; Suzuki et al., 2005), while a VEGFR3^{+}CD34^{+} progenitor population has been reported to be responsible for human postnatal lymphangiogenesis (Salven et al., 2003; Tan et al., 2014). However, to date, the progenitor population responsible for human lymphatic development at the embryonic stage has not been identified. My data shows, for the first time, the possible existence of a VEGFR3^{+}CD31^{+} human embryonic lymphatic progenitor population.

I now propose an updated model of the balloon self-organization system (Figure 4.18). In this model, Lymphgen-treatment first induces the formation of "clusters" containing DiI-Ac-LDL^{+}VE-cadherin^{+}CD31^{+}CD34^{+}KDR^{+}CD43^{-}VEGFR3^{-} (Endo^{+}VEGFR3^{-} for short) endothelial progenitors (blue) and Endo^{+}VEGFR3^{+} lymphatic progenitors (red). Together, these progenitors grow into "spheres" which may mimic venous vessels *in vivo*. The Endo^{+}VEGFR3^{+} lymphatic progenitor cells within those spheres form VEGFR3^{+} sprouts, which then grow into "balloons" that represent primitive lymphatic plexuses. The connections between balloons and spheres eventually constrict,

giving rise to stalks which may represent a remnant structure that mimics an incomplete separation between a primitive lymlphatic plexus and a venous vessel. In summary, my data strongly suggest that this spontaneous balloon self-organizing system recapitulates the complicated process of lymphatic morphogenesis *in vitro* without any external guidance, thus providing a valuable system for the study of early human lymphatic development while proving the feasibility of establishing a completely spontaneous self-organizing system of hESCs using small molecules as the initial stimuli.

Discussion

In this study I conducted a chemical screening designed based on a novel hypothesis for hESC differentiation induction, and successfully identified small molecules Lymphgen 1 and 2 as inducers of a spontaneous self-organizing 3-D organoid structure which was later demonstrated to recapitulate human lymphatic development. Free of growth-factor guidance, this method gives rise to the first truly spontaneous self-organizing organogenesis system, challenges the conventional thinking in the designing of organoid-induction protocols, and provides novel insights into our understandings of self-organizing processes or emergent behavior during organogenesis.

Furthermore, I successfully exploited this system to unveil the unguided emergence of a DiI-Ac-LDL^{+}VE-cadherin^{+}CD31^{+}CD34^{+}KDR^{+}CD43^{-} endothelial progenitor population during differentiation. I also identified a previously unknown VEGFR3^{+}CD31^{+} progenitor population which may be responsible for human lymphatic development at the embryonic stage. Last but not least, this organoid-formation system

provides a rare opportunity to visualize the truly spontaneous progression of human lymphatic morphogenesis *in vitro* for the first time.

Finally, this system provides a potential platform for the investigations of human diseases related to the lymphatic system, including hereditary diseases such as Milroy's disease (Butler et al., 2007), and many other common pathological conditions such as tumor migration and chronic inflammation (Schulte-Merker et al., 2011).

Materials and Methods

Cell culture

H9 and H1 hESC lines (WiCell Research Institute, Madison, WI) were maintained under a feeder condition or a feeder-independent condition. For the feeder condition (Thomson et al., 1998), primary mouse embryonic fibroblasts (MEFs) prepared from embryos of pregnant CF-1 mice (day 13.5 of gestation; Charles River) were cultured in Dulbecco's Modified Eagle Medium (DMEM) containing 10% FBS (Hyclone), 1% non-essential amino acids (NEAA; Invitrogen), and penicillin/streptomycin, and mitotically inactivated by gamma irradiation. H9 and H1 hESCs were cultured on irradiated MEFs in media containing DMEM/F12, 20% knockout serum replacement (KSR; Invitrogen), 4 ng/ml basic fibroblast growth factor (bFGF; Invitrogen), 1% NEAA, 1 mM glutamine, and 0.1 mM β-mercaptoethanol. For the feeder-independent condition, hESCs were cultured on Matrigel (BD Biosciences)-coated plates in mTeSR1 medium (StemCell Technologies) as described (Ludwig et al.,

2006). Experiments described in this study were conducted with H9 and H1 hESCs between passages 30 and 60.

Morphological screen

For the morphological screen, H9 and H1 hESCs were plated in Matrigel coated vessels as 2-D cultures in a basal medium composed of RPMI-1640 medium and 2% B-27 supplement (Invitrogen). Small molecules were individually applied to each culture, and incubated for 5 to 6 days before withdrawn. Cultures were then kept in the basal medium alone, without the addition of any growth factors, and let grown for 6 days or longer. During this prolonged incubation period, cultures were examined regularly for the emergence of 3-D organoid-like structures.

Balloon culture

H9 and H1 hESCs cultured on MEF feeders or Matrigel were dissociated by brief exposure to type IV collagenase (1 mg/ml; Invitrogen) or Dispase (Invitrogen), respectively, followed by mechanical dissociation into cell clumps. It is crucial to plate cells in clumps rather than single cells for the balloon formation. Cell clumps were seeded into Matrigel coated vessels in mTeSR1 medium at high density: generally at a 1:2 passaging ratio for MEF-cultured cells, and 1:1 for Matrigel-cultured cells. Cultures were then maintained in mTeSR1 medium with daily medium change until reaching approximately 90% confluency. Culture medium were then switched to a basal differentiation medium containing Advanced RPMI 1640 (Invitrogen), 2% B-27 supplement (Invitrogen), and 1% Glutamax (Invitrogen). Medium was changed every

other day. Lymphgen 1 and 2 were applied to the culture at 2.5 - 5 μM for the first 6 days. Balloons typically emerge at around day 10.

Selective adhesion assay

Cells from balloon cultures and DMSO control were trypsinized into single cells, counted using a hematocytometer, and seeded into Fibronectin (BD Biosciences) coated culture vessels in endothelial culture medium (ATCC) at a density of 1.5×10^4 cells/cm^2. Numbers of adherent cells and their endothelial identities were examined 3 days after seeding.

Matrigel assay

Cells from balloon cultures and DMSO control were trypsinized into single cells, counted using a hematocytometer, and encapsulated in Matrigel (kept cool to prevent gelation) at 1.0×10^6 cells/ml. Cultures were then kept in 37°C for 1 hour to allow formation of solid gels. Endothelial culture medium (ATCC) was added to each well. Medium was changed every other day. The emergences of vessel-like networks were monitored, and images were acquired using microscopy (Zeiss).

DiI-Ac-LDL staining

DiI-Ac-LDL were added to live cell culture at a 1:20 dilution and incubated at 37°C for 4 hr. Medium was then removed, and cells were washed with cell culture medium and visualized via fluorescent microscopy (Zeiss).

RNA extraction, reverse transcription, and quantitative-PCR

Total RNA were isolated using RNeasy mini kit (QIAGEN). cDNAs were synthesized from the purified RNAs using Reverse Transcription System (Promega). Quantitative-PCR was performed using QuantiTech SYBR Green PCR kit (QIAGEN). Signals were analyzed using the comparative C_T method, and *ACTB* gene was used as an internal control.

Flow cytometry

Single cell suspensions were acquired through Accutase (Invitrogen) treatment. Cells were stained using conjugated or unconjugated antibodies or UEA1 lectin for 30 min on ice. Cells stained by unconjugated antibodies were washed by ice cold FACS buffer (PBS supplemented with 1% BSA) and then stained using fluorescent-conjugated secondary antibodies for an extra 30 min. Cells were then washed and resuspended in ice cold FACS buffer and analyzed using a BD Biosciences LSR II flow cytometry analyzer and BD FACSDiva software.

Immunofluorescent staining

Cells were washed once with phosphate-buffered saline (PBS) and fixed by 4% paraformaldahyde (Santa Cruz) at room temperature for 15 min and then blocked with 5% donkey serum (Sigma) in PBS at room temperature for 1 hr. The samples were incubated with primary antibodies in PBST (PBS supplemented with 1% Triton) at 4°C overnight, washed three times by PBS, and then incubated with fluorescent-labeled secondary antibodies in PBST at room temperature for 1 hr. Finally, cells were incubated with DAPI

for 5 min, washed three times by PBS, and subjected to fluorescent microscopy analysis (Zeiss).

Statistical analysis

Statistical analyses were performed using Microsoft Excel. Student's t-test was used to compare two experimental groups, assuming unequal variances. Differences are considered significant when $p < 0.05$.

Figures

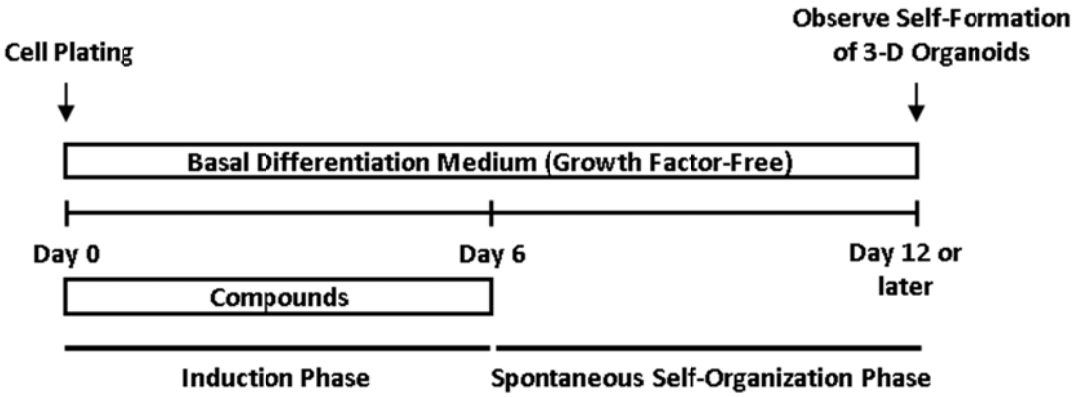


Figure 4.1. Schematic overview of the morphological screening.

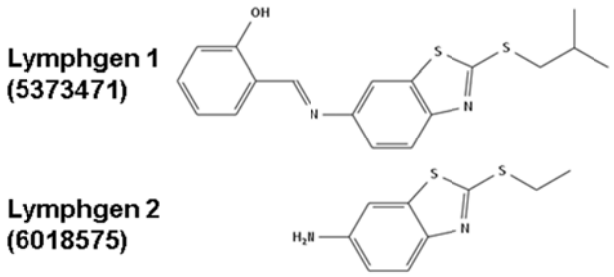


Figure 4.2. Chemical structures of 5373471 (Lymphgen 1) and 6018575 (Lymphgen 2).

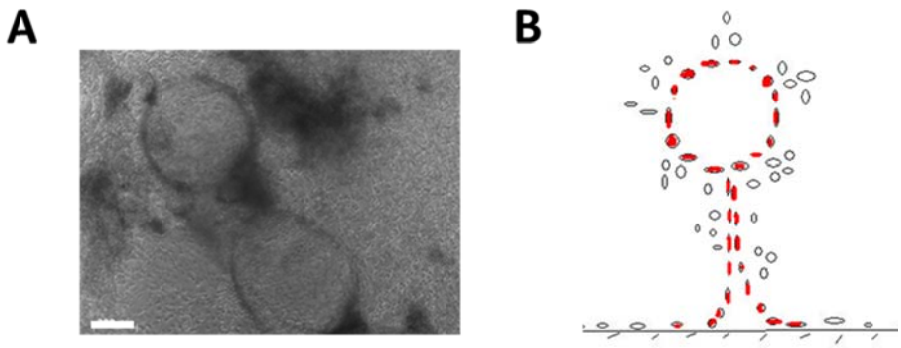


Figure 4.3. Image and illustration of the balloon-like structure. (**A**) Phase-contrast image of balloon-like structures. Scale bar: 200 μm. (**B**) Schematic drawing of a balloon-like structure. Red: the core part of a balloon showing a spherical structure and a stalk-like structure; Open circles: cells attached to the surface of the spherical structure and stalk.

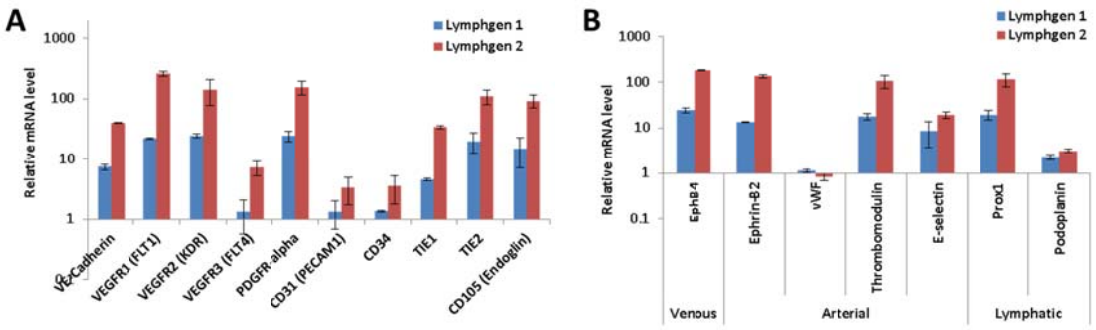


Figure 4.4. Lymphgens induce endothelial marker expression: qPCR. (**A**) Quantitative-PCR analysis of pan-endothelial markers in H9 hESCs treated with DMSO, 5373471 (Lymphgen 1, 5 μM), and 6018575 (Lymphgen 2, 5 μM) and cultured for 12 days. All values were normalized to the level (=1) of mRNA in cells treated with DMSO. Each bar represents mean±SD (error bars). *ACTB* (*β-actin*) was used as a loading control. (**B**) Quantitative-PCR analysis of venous, arterial, and lymphatic endothelial markers in H9 hESCs treated with DMSO, 5373471 (Lymphgen 1, 5 μM), and 6018575 (Lymphgen 2, 5 μM) and cultured for 12 days. All values were normalized to the level (=1) of mRNA in cells treated with DMSO. Each bar represents mean±SD (error bars). *ACTB* (*β-actin*) was used as a loading control.

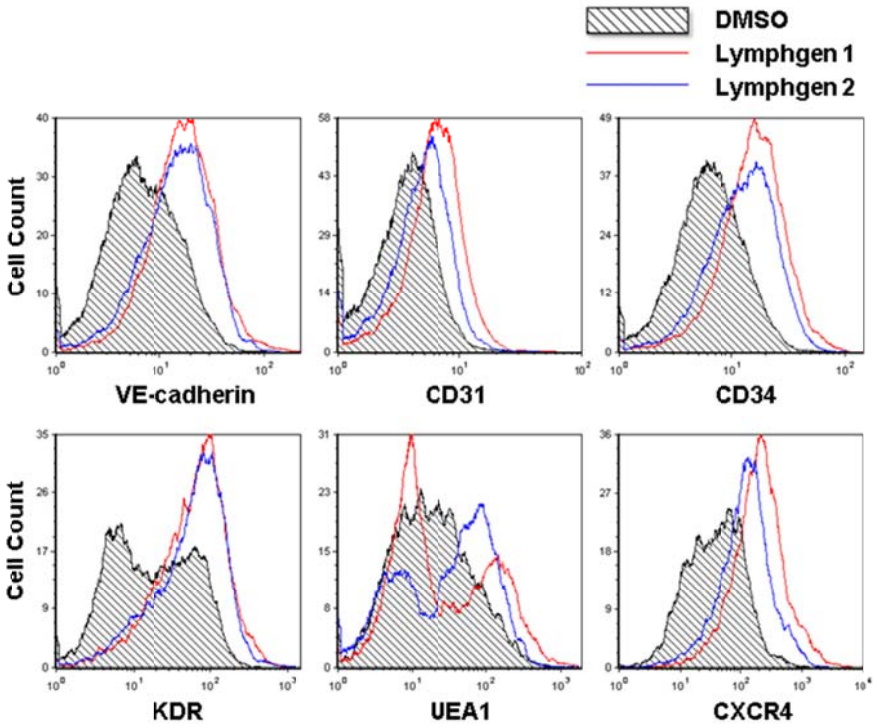


Figure 4.5. Lymphgens induce endothelial marker expression: FACS. FACS analysis of VE-cadherin, CD31, CD34, KDR, UEA1, and CXCR4 in H9 hESCs treated with DMSO, Lymphgen 1 (5 μM), and Lymphgen 2 (5 μM) and cultured for 12 days.

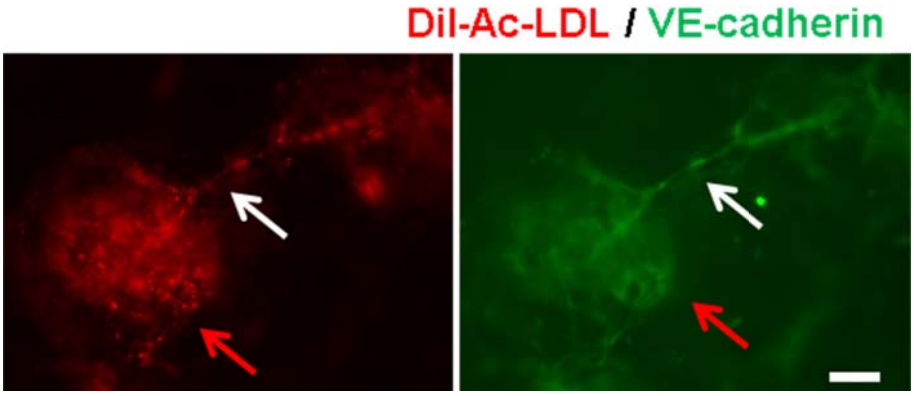


Figure 4.6. Lymphgens induce endothelial marker expression: immunofluorescent staining. Fluorescent images of balloons stained with DiI-Ac-LDL (red) and VE-cadherin (green) showing a spherical structure (red arrows) and a stalk (white arrows). Scale bar: 200 μm.

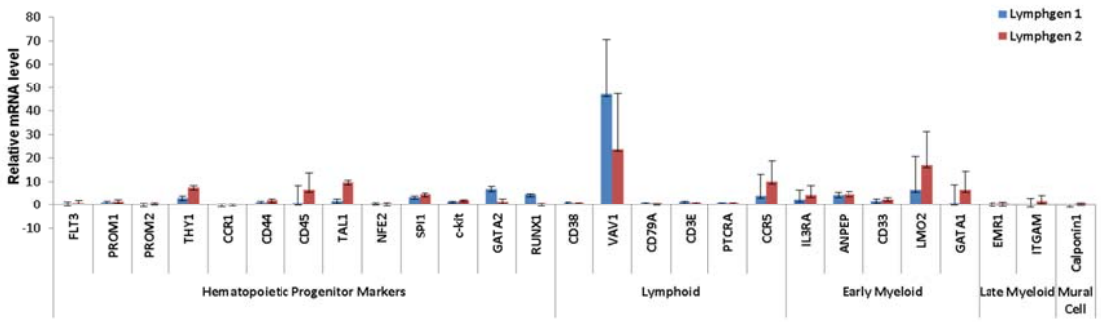


Figure 4.7. Lymphgens do not induce expressions of blood lineage- and mural cell-specific markers. Quantitative-PCR analysis of hematopoietic, lymphoid, myeloid, and mural cell markers in H9 hESCs treated with DMSO, Lymphgen 1 (5 μM), and Lymphgen 2 (5 μM) and cultured for 12 days. All values were normalized to the level (=1) of mRNA in cells treated with DMSO. Each bar represents mean ± SD (error bars). *ACTB* (*β-actin*) was used as a loading control.

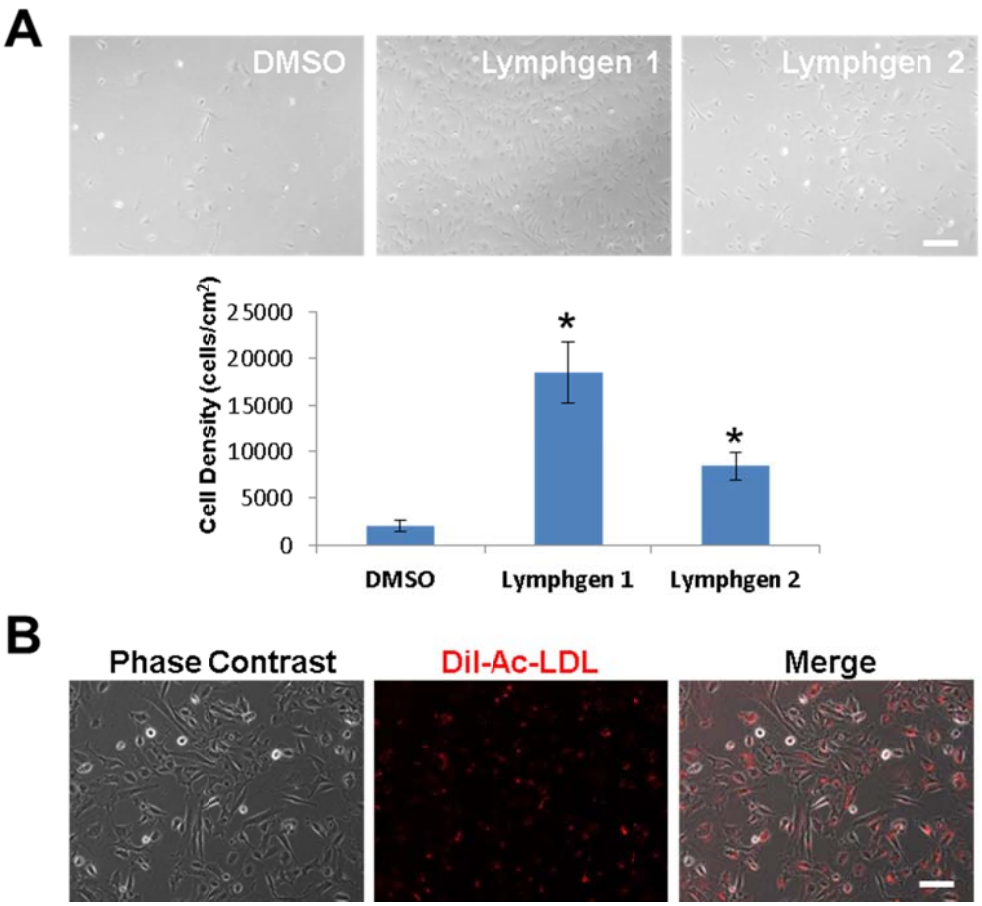


Figure 4.8. Functional tests confirm endothelial identities of Lymphgen-induced cultures: endothelial-selective re-plating. (A) Top: phase-contrast images of H9 hESCs treated by DMSO, Lymphgen 1 (5 μM), and Lymphgen 2 (5 μM) for 6 days, let grown till day 12, and then re-plated onto fibronectin coated plates and maintained in EC medium for 3 days. Scale bar: 200 μm. Bottom: quantification of cell densities in 3 experimental replicates, 3 days after re-plating. Each bar represents mean±SD (error bars). *: $p<0.01$. (**B**) Phase-contrast (left), fluorescent (middle), and merged (right) images of DiI-Ac-LDL in H9 hESCs treated by Lymphgen 1 (5 μM) for 6 days, cultured till day 12, and then replated onto fibronectin coated plates and maintained in EC medium for 3 days. Scale bar: 100 μm.

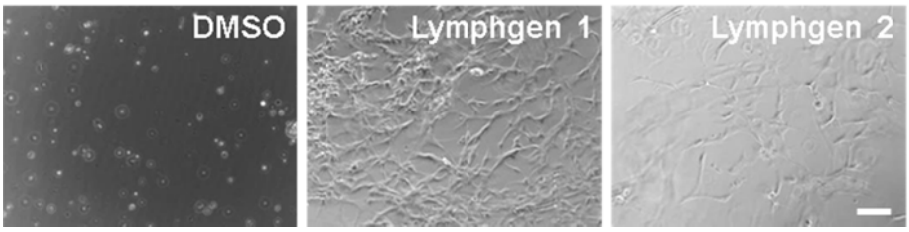


Figure 4.9. Functional tests confirm endothelial identities of Lymphgen-induced cultures: Matrigel assay. Phase-contrast images of H9 hESCs treated by DMSO, Lymphgen 1, and Lymphgen 2, cultured till day 12, and then encapsulated by Matrigel and maintained in EC medium for 5 days. Scale bar: 200 μm.

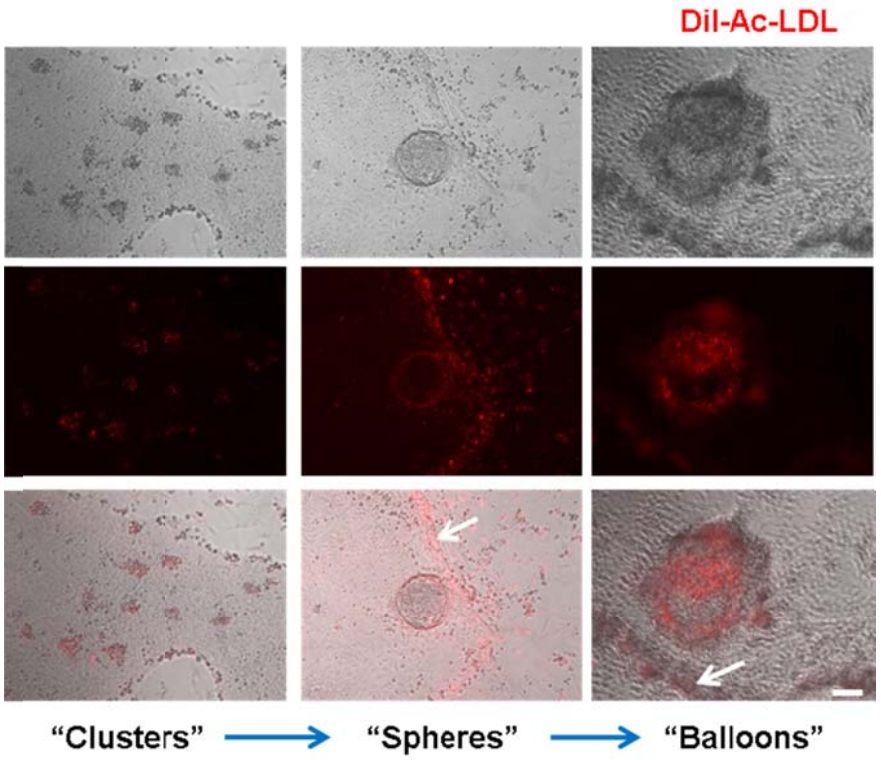


Figure 4.10. A working model of balloon self-organization: DiI-Ac-LDL staining. Fluorescent images of DiI-Ac-LDL (red) in H9 hESCs treated with Lymphgen 2 (5 μM) and cultured for 13 days, showing DiI-Ac-LDL$^+$ clusters (left), spheres (middle), and balloons (right). White arrows: irregularly shaped DiI-Ac-LDL$^+$ cell aggregates. Scale bar: 100 μm.

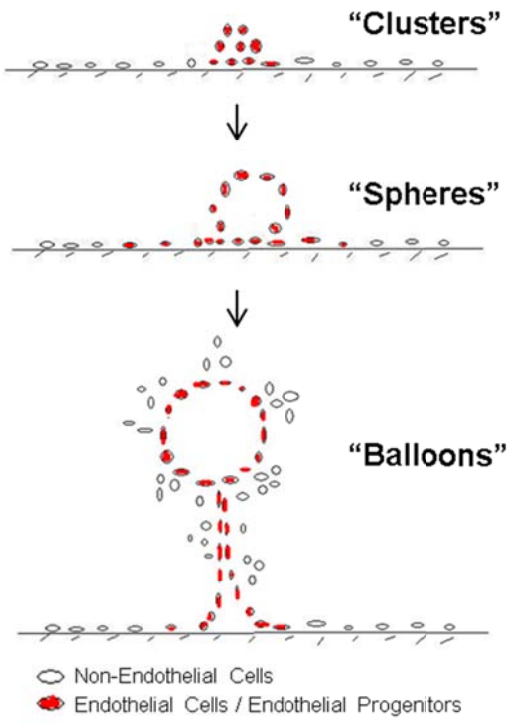


Figure 4.11. A working model of balloon self-organization: schematic drawing. Schematic drawing of a working model of balloon self-organization.

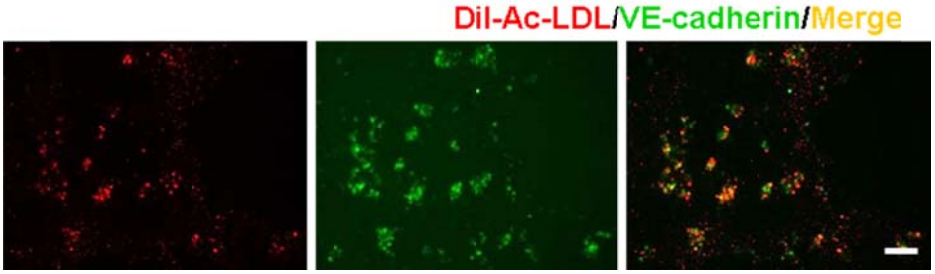


Figure 4.12. Identification of a DiI-Ac-LDL^{+}VE-cadherin^{+}CD31^{+}CD34^{+}KDR^{+}CD43^{-} endothelial progenitor population: immunofluorescent staining. Fluorescent images of DiI-Ac-LDL (red) and VE-cadherin (green) in H9 hESCs treated with Lymphgen 2 (5 μM) and cultured for 13 days showing DiI-Ac-LDL^{+}VE-cadherin^{+} clusters. Scale bar: 100 μm.

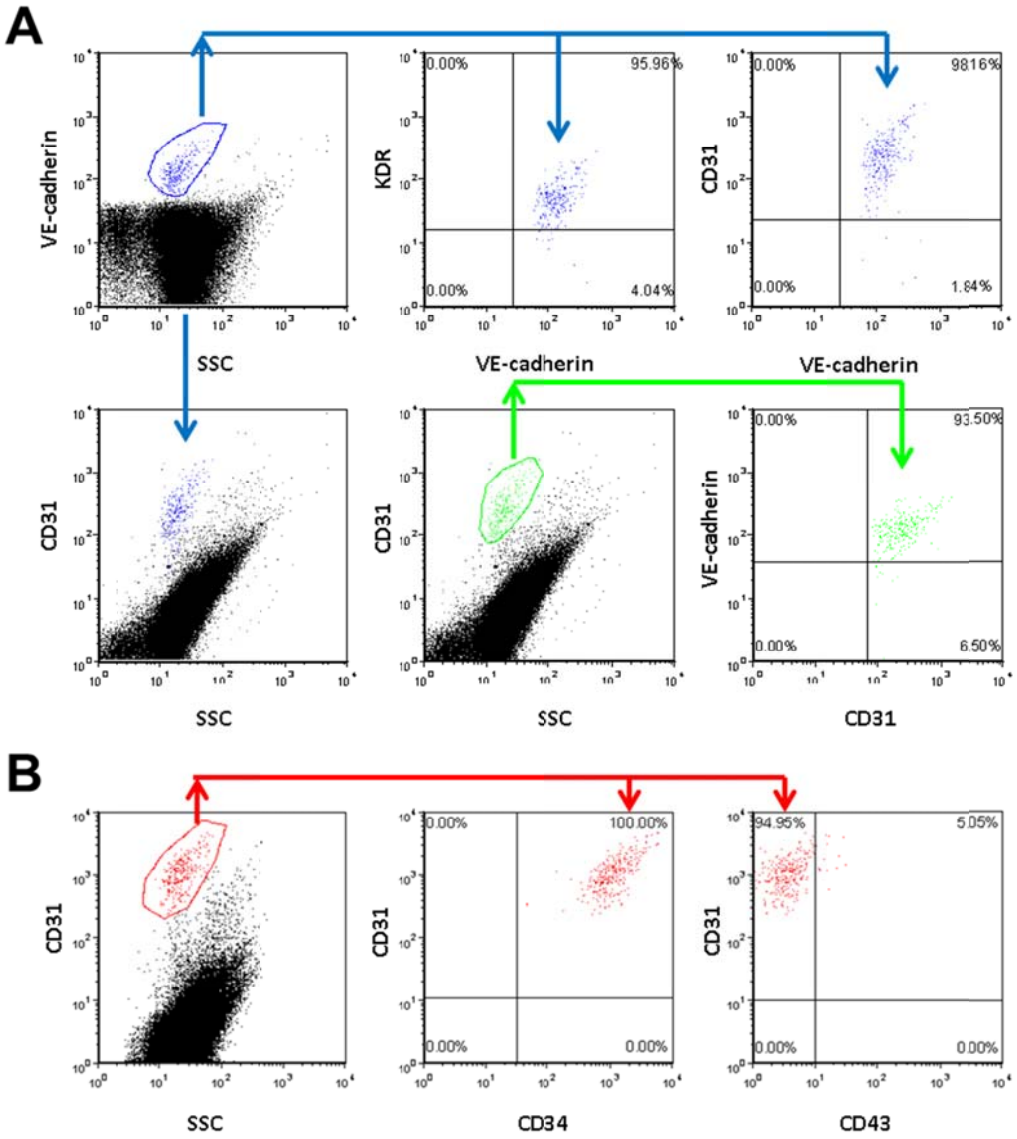


Figure 4.13. Identification of a DiI-Ac-LDL^{+}VE-cadherin^{+}CD31^{+}CD34^{+}KDR^{+}CD43^{-} endothelial progenitor population: FACS. (**A**) FACS analysis of VE-cadherin, CD31, and KDR in H9 hESCs treated with Lymphgen 1 (5 μM) and cultured for 13 days, showing a distinct VE-cadherin^{+}CD31^{+}KDR^{+} population by gating VE-cadherin (top panel). CD31^{+} cells also appear as a distinct population which is also VE-cadherin^{+} (bottom panel). (**B**) FACS analysis of CD31, CD34, and CD43 in H9 hESCs treated with Lymphgen 1 (5 μM) and cultured for 12 days, showing a distinct CD31^{+}CD34^{+}CD43^{-} population.

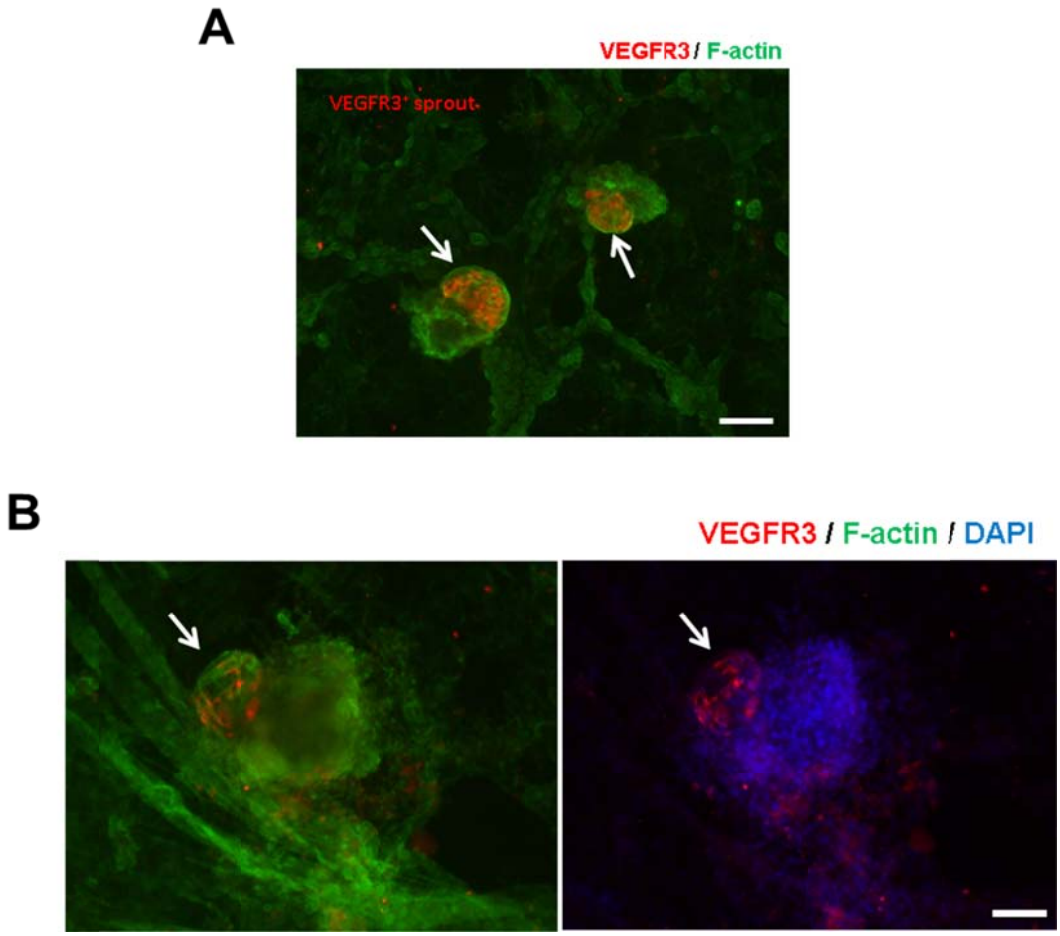


Figure 4.14. Balloon self-organization may recapitulate lymphatic vascular morphogenesis: VEGFR3$^+$ sprouts. (**A**) Immunofluorescent staining images of VEGFR3 (red) and F-actin (green) in two "spheres" formed in a H9 hESC culture treated with Lymphgen 2 (5 μM) and cultured for 29 days, showing VEGFR3$^+$ sprouts. Scale bar: 100 μm. (**B**) Immunofluorescent staining images of VEGFR3 (red), F-actin (green), and DAPI (blue) in a "sphere" formed in a H9 hESC culture treated with Lymphgen 2 (5 μM) and cultured for 29 days, showing a budding VEGFR3$^+$ sprout (white arrows). Scale bar: 100 μm.

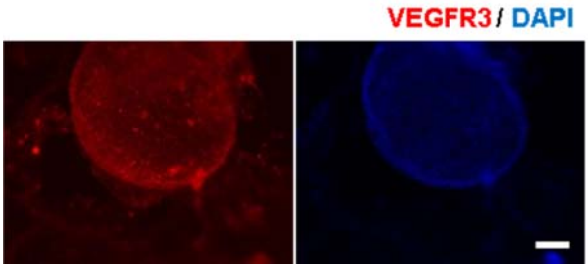


Figure 4.15. Balloon self-organization may recapitulate lymphatic vascular morphogenesis: VEGFR3$^+$ balloons. Immunofluorescent staining of VEGFR3 (red) and DAPI (blue) in a balloon formed in a H9 hESC culture treated with Lymphgen 2 (5 μM) and cultured for 29 days. Scale bar: 100 μm.

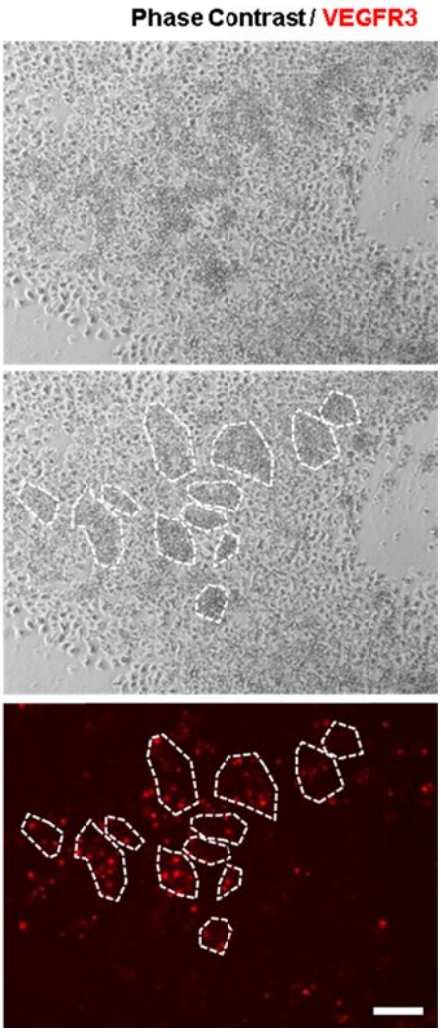


Figure 4.16. Balloon self-organization may recapitulate lymphatic vascular morphogenesis: VEGFR3^{+} progenitors in "clusters". Phase-contrast images (top and middle) and an immunofluorescent staining image (bottom) of VEGFR3 (red) showing "clusters" (enclosed by white dotted lines) formed in a H9 hESC culture treated with Lymphgen 2 (5 μM) and cultured for 13 days. Scale bar: 100 μm.

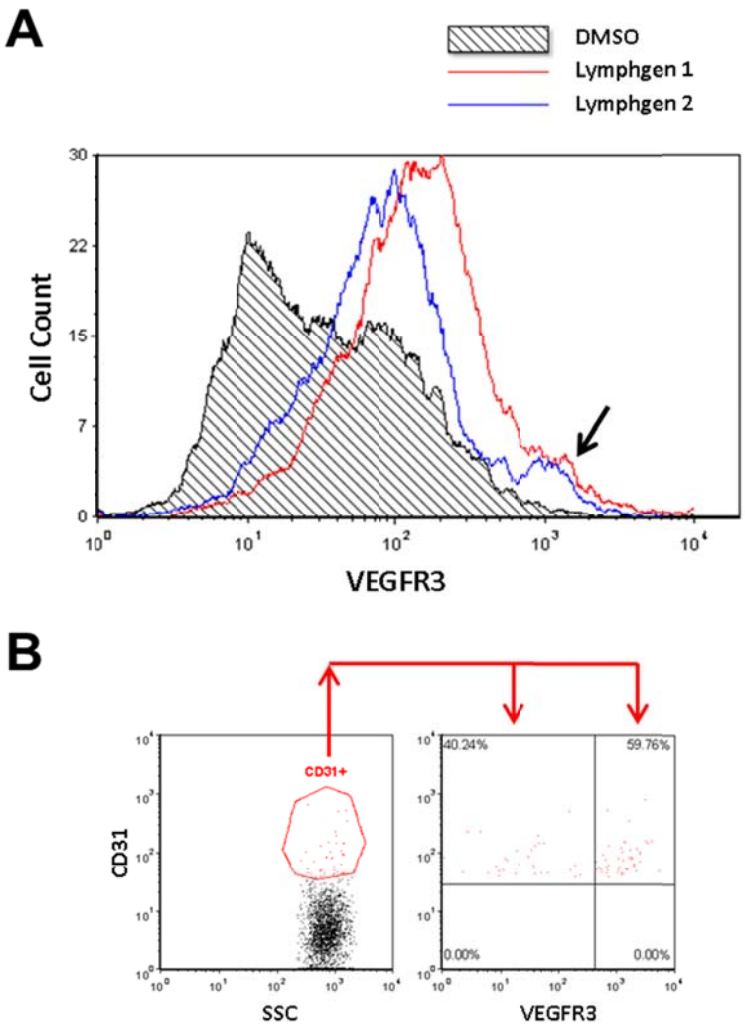


Figure 4.17. Balloon self-organization may recapitulate lymphatic vascular morphogenesis: VEGFR3^{+}CD31^{+} progenitors. (**A**) FACS analysis of VEGFR3 in H9 hESCs treated DMSO, Lymphgen 1 (5 μM), and Lymphgen 2 (5 μM) and cultured for 12 days, showing a minor peak of VEGFR3^{+} population (black arrow) in Lymphgen 1 and 2 treated cells. (**B**) FACS analysis of VEGFR3 and CD31 in H9 hESCs treated with Lymphgen 1 (5 μM) and cultured for 12 days showing a VEGFR3^{+}CD31^{+} population.

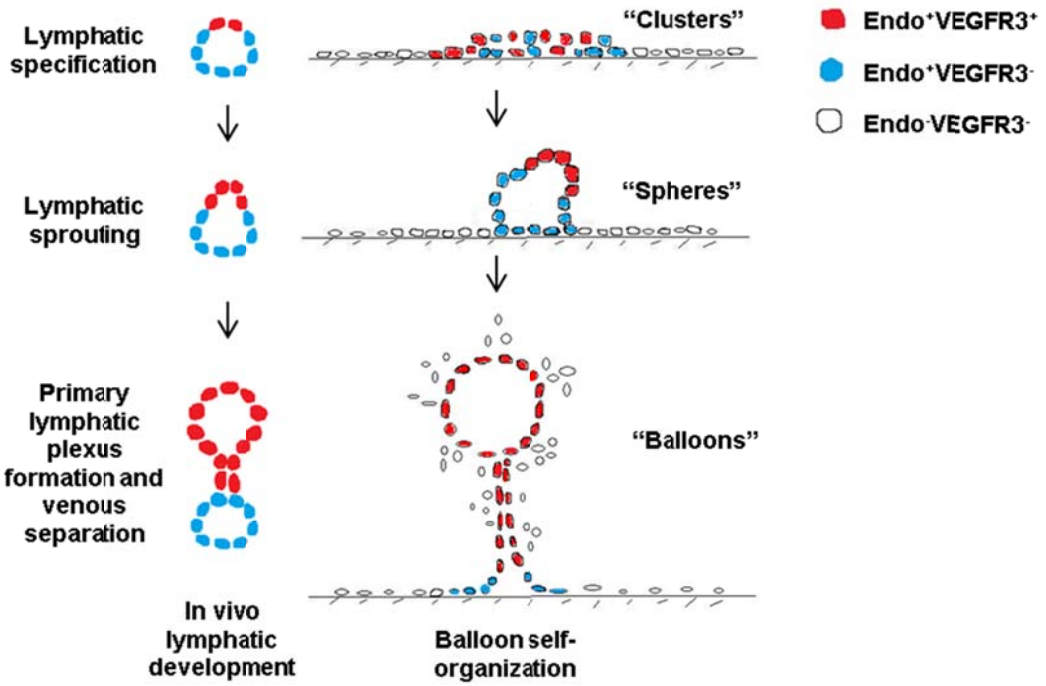


Figure 4.18. Balloon self-organization recapitulates the *in vivo* lymphatic morphogenesis. Schematic drawings comparing *in vivo* lymphatic vessel development and balloon self-organization.

Tables

Antibody	Source	Catalog Number	Dilution
VE-cadherin	R&D Systems	AF938	1:500 (IF) / 1:50 (FACS)
VEGFR3	Millipore	MAB3757	1:50 (IF) / 1:50 (FACS)
CD31-PE	BD Biosciences	560983	1:5 (FACS)
CD31-FITC	BD Biosciences	555445	1:5 (FACS)
VE-cadherin-Alexa647	BD Biosciences	561567	1:20 (FACS)
KDR-PE	BD Biosciences	560494	1:5 (FACS)
CD34-APC	BD Biosciences	560940	1:5 (FACS)
CD43-FITC	BD Biosciences	560978	1:5 (FACS)
CXCR4	R&D Systems	MAB172	1:1000 (FACS)
CD34	Millipore	CBL496	1:50 (FACS)
KDR	Millipore	05-554	1:50 (FACS)
Donkey anti-mouse, Alexa 488	Invitrogen	A-21202	1:200 (FACS)
Donkey anti-goat, Alexa 488	Invitrogen	A-11055	1:200 (FACS)
Donkey anti-mouse, Alexa 594	Invitrogen	A-21203	1:200 (FACS)
UEA1-Fluorescein	Vector Labs	FL-1061	1:1000 (FACS)
DiI-Ac-LDL	Biomedical Technologies	BT-902	1:20 (IF)
Phalloidin, Alexa 488	Invitrogen	A12379	1:1000 (IF)

(IF: immunofluorescence)

Table 4.1. Antibodies and fluorescent dyes

Primer Name	Forward Primer (5'-3')	Reverse Primer (5'-3')
ACTB	agagctacgagctgcctgac	cgtggatgccacaggact
VE-cadherin	gttcacgcatcggttgttcaa	cgcttccaccacgatctcata
VEGFR1 (FLT1)	gtcacagaagaggatgaaggtgtc	cacagtccggcacgtaggtgatt
VEGFR2 (KDR)	tttttgcccttgttctgtcc	tcattgttcccagcatttca
VEGFR3 (FLT4)	agacaagaaagcggcttcag	cctcccttgggagtcagg
PDGFRA	ccacctgagtgagattgtgg	tcttcaggaagtccaggtgaa
CD31 (PECAM1)	tctatgacctcgccctccacaaa	gaacggtgtcttcaggttggtatttca
CD34	gcgctttgcttgctgagt	gggtagcagtaccgttgttgt
TIE1	agaacctagcctccaagatt	actgtagttcagggactcaa
TIE2	tccaaggatgtctctgctctc	ttggggtcatcctcggtat
CD105 (Endoglin)	ttcctggagttcccaacg	aggtgccattttgcttgg
EphB4	cgcacctacgaagtgtgtga	gtccgcatcgctctcatagta
Ephrin-B2	tatgcagaactgcgatttccaa	tgggtatagtaccagtccttgtc
vWF	agtgcagacccaacttcacc	gtggggacactcttttgcac
Thrombomodulin	aattgggagcttgggaatg	tgaggacctgattaaggctagg
E-selectin	accagcccaggttgaatg	ggttggacaaggctgtgc
PROX1	gagcctccgtggaactca	tgggcacagctcaagaatc
Podoplanin	aaatgtcgggaaggtactcg	agggcacagagtcagaaacg
FLT3	tggaatttctggaatttaagtcg	tttcccgtgggtgacaag
PROM1	tccacagaaatttacctacattgg	cagcagagagcagatgacca
PROM2	tgcagctcaacgactcctac	actcctgccgtagcttgttg
THY1	aggacgagggcacctacac	gccctcacacttgaccagtt
CCR1	tctgactcttggcacagcat	gccaccattacattccctctc
CD44	caacaacacaaatggctggt	ctgaggtgtctgtctctttcatct
CD45	ccaatgcaaaactcaacccta	cctctctcctgggacatctg
TAL1	accctagccctctccctcta	ggccacaggctatctctcct
NFE2	gggagcctcatctctctcct	agcccagatggctctagaaa
SPI1	caggggatctgaccgactc	aggtcttctgatggctgagg
c-kit	aagagatgtgactcccgccat	aggaaagccatgcccttg
GATA2	aaggctcgttcctgttcaga	ggcattgcacaggtagtgg
RUNX1	ccctaggggatgttccagat	tgaagcttttccctcttcca
CD38	cagcaacaaccctgtttcagt	ccattgagcatcacatggac
VAV1	agcagacggaggagaagtaca	caggggcttcaagaaatgc
CD79A	caagaaccgaatcatcacagc	cgttctgccatcgtttcc
CD3E	caaggccaagcctgtgac	tcatagtctgggttgggaaca
PTCRA	ccagccctacccacaggt	atcaaggaccaggcagacc
CCR5	aaccaggcgagagacttgtg	gatccaactcaaattccttctca
IL3RA	tcccgatggtcctccttt	tgattggtgggtttggatct

Table 4.2. qPCR primers

Table 4.2. (Cont.)

ANPEP	catccatcagagatggcagac	tgctgaagagatcgttctgg
CD33	caggaatgacacccacccta	tcagtggggccatgtaactt
LMO2	ttcaacctgaacttgcagtagc	gcttccctctgtctctggttt
GATA1	cactgagcttgccacatcc	atggagcctctggggatta
EMR1	gccagtttgcacagttatagtga	gacttcggcattaacactgga
ITGAM	ggcatccgcaaagtggta	ggatcttaaaggcattctttcg
Calponin1	ctgtcagccgaggttaagaac	gaggccgtccatgaagttgtt

REFERENCES

Ahn, S.G., Kim, S.A., Yoon, J.H., and Vacratsis, P. (2005). Heat-shock cognate 70 is required for the activation of heat-shock factor 1 in mammalian cells. Biochem J *392*, 145-152.

Andrews, P.W. (1984). Retinoic acid induces neuronal differentiation of a cloned human embryonal carcinoma cell line in vitro. Dev Biol *103*, 285-293.

Andrews, P.W., Damjanov, I., Berends, J., Kumpf, S., Zappavigna, V., Mavilio, F., and Sampath, K. (1994). Inhibition of Proliferation and Induction of Differentiation of Pluripotent Human Embryonal Carcinoma-Cells by Osteogenic Protein-1 (or Bone Morphogenetic Protein-7). Lab Invest *71*, 243-251.

Andrews, P.W., Damjanov, I., Simon, D., Banting, G.S., Carlin, C., Dracopoli, N.C., and Fogh, J. (1984). Pluripotent Embryonal Carcinoma Clones Derived from the Human Teratocarcinoma Cell-Line Tera-2 - Differentiation Invivo and Invitro. Lab Invest *50*, 147-162.

Andrews, P.W., Gonczol, E., Plotkin, S.A., Dignazio, M., and Oosterhuis, J.W. (1986). Differentiation of TERA-2 human embryonal carcinoma cells into neurons and HCMV permissive cells. Induction by agents other than retinoic acid. Differentiation *31*, 119-126.

Andrews, P.W., Nudelman, E., Hakomori, S.I., and Fenderson, B.A. (1990). Different Patterns of Glycolipid Antigens Are Expressed Following Differentiation of Tera-2 Human Embryonal Carcinoma-Cells Induced by Retinoic Acid, Hexamethylene Bisacetamide (Hmba) or Bromodeoxyuridine (Budr). Differentiation *43*, 131-138.

Antonica, F., Kasprzyk, D.F., Opitz, R., Iacovino, M., Liao, X.H., Dumitrescu, A.M., Refetoff, S., Peremans, K., Manto, M., Kyba, M.*, et al.* (2012). Generation of functional thyroid from embryonic stem cells. Nature *491*, 66-71.

Barbaric, I., Gokhale, P.J., Jones, M., Glen, A., Baker, D., and Andrews, P.W. (2010). Novel regulators of stem cell fates identified by a multivariate phenotype screen of small compounds on human embryonic stem cell colonies. Stem Cell Res *5*, 104-119.

Basson, M.A., Watson-Johnson, J., Shakya, R., Akbulut, S., Hyink, D., Costantini, F.D., Wilson, P.D., Mason, I.J., and Licht, J.D. (2006). Branching morphogenesis of the ureteric epithelium during kidney development is coordinated by the opposing functions of GDNF and Sprouty1. Dev Biol *299*, 466-477.

Bernardo, A.S., Faial, T., Gardner, L., Niakan, K.K., Ortmann, D., Senner, C.E., Callery, E.M., Trotter, M.W., Hemberger, M., Smith, J.C.*, et al.* (2011). BRACHYURY and CDX2 mediate BMP-induced differentiation of human and mouse pluripotent stem cells into embryonic and extraembryonic lineages. Cell Stem Cell *9*, 144-155.

Borowiak, M., Maehr, R., Chen, S., Chen, A.E., Tang, W., Fox, J.L., Schreiber, S.L., and Melton, D.A. (2009a). Small Molecules Efficiently Direct Endodermal Differentiation of Mouse and Human Embryonic Stem Cells. Cell Stem Cell *4*, 348-358.

Borowiak, M., Maehr, R., Chen, S., Chen, A.E., Tang, W., Fox, J.L., Schreiber, S.L., and Melton, D.A. (2009b). Small molecules efficiently direct endodermal differentiation of mouse and human embryonic stem cells. Cell Stem Cell *4*, 348-358.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., *et al.* (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell *122*, 947-956.

Butler, M.G., Dagenais, S.L., Rockson, S.G., and Glover, T.W. (2007). A novel VEGFR3 mutation causes Milroy disease. Am J Med Genet A *143A*, 1212-1217.

Cai, J., Zhao, Y., Liu, Y.X., Ye, F., Song, Z.H., Qin, H., Meng, S., Chen, Y.Z., Zhou, R.D., Song, X.J., *et al.* (2007). Directed differentiation of human embryonic stem cells into functional hepatic cells. Hepatology *45*, 1229-1239.

Cao, L., Gibson, J.D., Miyamoto, S., Sail, V., Verma, R., Rosenberg, D.W., Nelson, C.E., and Giardina, C. (2011). Intestinal lineage commitment of embryonic stem cells. Differentiation *81*, 1-10.

Caron, N.J., Gage, B.K., O'Connor, M.D., Eaves, C.J., Kieffer, T.J., and Piret, J.M. (2013). A Human Embryonic Stem Cell Line Adapted for High Throughput Screening. Biotechnol Bioeng *110*, 2706-2716.

Carter, C., Pan, S., Zouhar, J., Avila, E.L., Girke, T., and Raikhel, N.V. (2004). The vegetative vacuole proteome of Arabidopsis thaliana reveals predicted and unexpected proteins. Plant Cell *16*, 3285-3303.

Catena, R., Tiveron, C., Ronchi, A., Porta, S., Ferri, A., Tatangelo, L., Cavallaro, M., Favaro, R., Ottolenghi, S., Reinbold, R., *et al.* (2004). Conserved POU binding DNA sites in the Sox2 upstream enhancer regulate gene expression in embryonic and neural stem cells. Journal of Biological Chemistry *279*, 41846-41857.

Chambers, S.M., Fasano, C.A., Papapetrou, E.P., Tomishima, M., Sadelain, M., and Studer, L. (2009). Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. Nat Biotechnol *27*, 275-280.

Chan, Y.S., Goke, J., Ng, J.H., Lu, X.Y., Gonzales, K.A.U., Tan, C.P., Tng, W.Q., Hong, Z.Z., Lim, Y.S., and Ng, H.H. (2013). Induction of a Human Pluripotent State with Distinct Regulatory Circuitry that Resembles Preimplantation Epiblast. Cell Stem Cell *13*, 663-675.

Chen, S., Borowiak, M., Fox, J.L., Maehr, R., Osafune, K., Davidow, L., Lam, K., Peng, L.F., Schreiber, S.L., Rubin, L.L., *et al.* (2009). A small molecule that directs differentiation of human ESCs into the pancreatic lineage. Nat Chem Biol *5*, 258-265.

Chen, S., Do, J.T., Zhang, Q., Yao, S., Yan, F., Peters, E.C., Scholer, H.R., Schultz, P.G., and Ding, S. (2006a). Self-renewal of embryonic stem cells by a small molecule. Proc Natl Acad Sci U S A *103*, 17266-17271.

Chen, S.B., Do, J.T., Zhang, Q.S., Yao, S.Y., Yan, F., Peters, E.C., Scholer, H.R., Schultz, P.G., and Ding, S. (2006b). Self-renewal of embryonic stem cells by a small molecule. P Natl Acad Sci USA *103*, 17266-17271.

Chen, Y.S., Pelekanos, R.A., Ellis, R.L., Horne, R., Wolvetang, E.J., and Fisk, N.M. (2012). Small Molecule Mesengenic Induction of Human Induced Pluripotent Stem Cells to Generate Mesenchymal Stem/Stromal Cells. Stem Cell Transl Med *1*, 83-95.

Chetty, S., Pagliuca, F.W., Honore, C., Kweudjeu, A., Rezania, A., and Melton, D.A. (2013). A simple tool to improve pluripotent stem cell differentiation. Nat Methods *10*, 553-556.

Chew, J.L., Loh, Y.H., Zhang, W.S., Chen, X., Tam, W.L., Yeap, L.S., Li, P., Ang, Y.S., Lim, B., Robson, P., *et al.* (2005). Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. Molecular and Cellular Biology *25*, 6031-6046.

Cho, H.J., Gee, H.Y., Baek, K.H., Ko, S.K., Park, J.M., Lee, H., Kim, N.D., Lee, M.G., and Shin, I. (2011). A Small Molecule That Binds to an ATPase Domain of Hsc70 Promotes Membrane Trafficking of Mutant Cystic Fibrosis Transmembrane Conductance Regulator. J Am Chem Soc *133*, 20267-20276.

Chou, Y.F., Chen, H.H., Eijpe, M., Yabuuchi, A., Chenoweth, J.G., Tesar, P., Lu, J., McKay, R.D.G., and Geijsen, N. (2008). The Growth Factor Environment Defines Distinct Pluripotent Ground States in Novel Blastocyst-Derived Stem Cells. Cell *135*, 449-461.

Christenson, L.K., and Stouffer, R.L. (1996). Isolation and culture of microvascular endothelial cells from the primate corpus luteum. Biol Reprod *55*, 1397-1404.

Chubanov, V., Gudermann, T., and Schlingmann, K.P. (2005a). Essential role for TRPM6 in epithelial magnesium transport and body magnesium homeostasis. Pflug Arch Eur J Phy *451*, 228-234.

Chubanov, V., Schlingmann, K.P., Waring, J., Heinzinger, J., Kaske, S., Waidegger, S., Schnitzler, M.M.Y., and Gudermann, T. (2007). Hypomagnesemia with secondary hypocalcemia due to a missense mutation in the putative pore-forming region of TRPM6. Journal of Biological Chemistry *282*, 7656-7667.

Chubanov, V., Schlingmann, K.P., Waring, J., Schnitzler, M.M.Y., Waldegger, S., and Gudermann, T. (2006). Dominant-negative effect of a novel missense mutation in the human TRPM6 gene associated with hypomagnesemia with secondary hypocalcemia. N-S Arch Pharmacol *372*, 60-61.

Chubanov, V., Schnitzler, M.M.Y., Waring, J., Plank, A., and Gudermann, T. (2005b). Emerging roles of TRPM6/TRPM7 channel kinase signal transduction complexes. N-S Arch Pharmacol *371*, 334-341.

Chubanov, V., Waldegger, S., Schnitzler, M.M., Vitzthum, H., Sassen, M.C., Seyberth, H.W., Konrad, M., and Gudermann, T. (2004). Disruption of TRPM6/TRPM7 complex formation by a mutation in the TRPM6 gene causes hypomagnesemia with secondary hypocalcemia. P Natl Acad Sci USA *101*, 2894-2899.

D'Amour, K.A., Agulnick, A.D., Eliazer, S., Kelly, O.G., Kroon, E., and Baetge, E.E. (2005). Efficient differentiation of human embryonic stem cells to definitive endoderm. Nat Biotechnol *23*, 1534-1541.

D'Amour, K.A., Bang, A.G., Eliazer, S., Kelly, O.G., Agulnick, A.D., Smart, N.G., Moorman, M.A., Kroon, E., Carpenter, M.K., and Baetge, E.E. (2006). Production of pancreatic hormone-expressing endocrine cells from human embryonic stem cells. Nature Biotechnology *24*, 1392-1401.

Desbordes, S.C., Placantonakis, D.G., Ciro, A., Socci, N.D., Lee, G., Djaballah, H., and Studer, L. (2008). High-throughput screening assay for the identification of compounds regulating self-renewal and differentiation in human embryonic stem cells. Cell Stem Cell *2*, 602-612.

DeZwaan, D.C., and Freeman, B.C. (2010). HSP90 manages the ends. Trends Biochem Sci *35*, 384-391.

Ding, S., Wu, T.Y.H., Brinker, A., Peters, E.C., Hur, W., Gray, N.S., and Schultz, P.G. (2003). Synthetic small molecules that control stem cell fate. P Natl Acad Sci USA *100*, 7632-7637.

Draper, J.S., Pigott, C., Thomson, J.A., and Andrews, P.W. (2002). Surface antigens of human embryonic stem cells: changes upon differentiation in culture. J Anat *200*, 249-258.

Eiraku, M., Takata, N., Ishibashi, H., Kawada, M., Sakakura, E., Okuda, S., Sekiguchi, K., Adachi, T., and Sasai, Y. (2011). Self-organizing optic-cup morphogenesis in three-dimensional culture. Nature *472*, 51-56.

Eiraku, M., Watanabe, K., Matsuo-Takasaki, M., Kawada, M., Yonemura, S., Matsumura, M., Wataya, T., Nishiyama, A., Muguruma, K., and Sasai, Y. (2008). Self-organized formation of polarized cortical tissues from ESCs and its active manipulation by extrinsic signals. Cell Stem Cell *3*, 519-532.

Ellis, R.J. (2007). Protein misassembly: Macromolecular crowding and molecular chaperones. Adv Exp Med Biol *594*, 1-13.

Esteban, M.A., Wang, T., Qin, B., Yang, J., Qin, D., Cai, J., Li, W., Weng, Z., Chen, J., Ni, S., *et al.* (2009). Vitamin C enhances the generation of mouse and human induced pluripotent stem cells. Cell Stem Cell *6*, 71-79.

Evseenko, D., Zhu, Y., Schenke-Layland, K., Kuo, J., Latour, B., Ge, S., Scholes, J., Dravid, G., Li, X., MacLellan, W.R., *et al.* (2010). Mapping the first stages of mesoderm commitment during differentiation of human embryonic stem cells. Proc Natl Acad Sci U S A *107*, 13742-13747.

Fathi, A., Pakzad, M., Taei, A., Brink, T.C., Pirhaji, L., Ruiz, G., Sharif Tabe Bordbar, M., Gourabi, H., Adjaye, J., Baharvand, H., *et al.* (2009). Comparative proteome and transcriptome analyses of embryonic stem cells during embryoid body-based differentiation. Proteomics *9*, 4859-4870.

Freeman, B.C., Myers, M.P., Schumacher, R., and Morimoto, R.I. (1995). Identification of a Regulatory Motif in Hsp70 That Affects Atpase Activity, Substrate-Binding and Interaction with Hdj-1. Embo J *14*, 2281-2292.

Freeman, B.C., and Yamamoto, K.R. (2002). Disassembly of transcriptional regulatory complexes by molecular chaperones. Science *296*, 2232-2235.

Gerhardt, H. (2008). VEGF and endothelial guidance in angiogenic sprouting. Organogenesis *4*, 241-246.

Gerrard, L., Zhao, D., Clark, A.J., and Cui, W. (2005). Stably transfected human embryonic stem cell clones express OCT4-specific green fluorescent protein and maintain self-renewal and pluripotency. Stem Cells *23*, 124-133.

Ghabrial, A.S., and Krasnow, M.A. (2006). Social interactions among epithelial cells during tracheal branching morphogenesis. Nature *441*, 746-749.

Gifford, C.A., Ziller, M.J., Gu, H.C., Trapnell, C., Donaghey, J., Tsankov, A., Shalek, A.K., Kelley, D.R., Shishkin, A.A., Issner, R., *et al.* (2013). Transcriptional and Epigenetic Dynamics during Specification of Human Embryonic Stem Cells. Cell *153*, 1149-1163.

Goloudina, A.R., Demidov, O.N., and Garrido, C. (2012). Inhibition of HSP70: a challenging anti-cancer strategy. Cancer Lett *325*, 117-124.

Gonzalez, R., Lee, J.W., and Schultz, P.G. (2011a). Stepwise Chemically Induced Cardiomyocyte Specification of Human Embryonic Stem Cells. Angew Chem Int Edit *50*, 11181-11185.

Gonzalez, R., Lee, J.W., Snyder, E.Y., and Schultz, P.G. (2011b). Dorsomorphin Promotes Human Embryonic Stem Cell Self-Renewal. Angew Chem Int Edit *50*, 3439-3441.

Gupta, S.K., Lysko, P.G., Pillarisetti, K., Ohlstein, E., and Stadel, J.M. (1998). Chemokine receptors in human endothelial cells. Functional expression of CXCR4 and its transcriptional regulation by inflammatory cytokines. J Biol Chem *273*, 4282-4287.

Hanna, J., Cheng, A.W., Saha, K., Kim, J., Lengner, C.J., Soldner, F., Cassady, J.P., Muffat, J., Carey, B.W., and Jaenisch, R. (2010). Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. P Natl Acad Sci USA *107*, 9222-9227.

Hu, Q., and Rosenfeld, M.G. (2012). Epigenetic regulation of human embryonic stem cells. Front Genet *3*, 238.

Huangfu, D., Osafune, K., Maehr, R., Guo, W., Eijkelenboom, A., Chen, S., Muhlestein, W., and Melton, D.A. (2008). Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. Nat Biotechnol *26*, 1269-1275.

Huigens, R.W., Morrison, K.C., Hicklin, R.W., Flood, T.A., Richter, M.F., and Hergenrother, P.J. (2013). A ring-distortion strategy to construct stereochemically complex and structurally diverse compounds from natural products. Nat Chem *5*, 195-202.

Itskovitz-Eldor, J., Schuldiner, M., Karsenti, D., Eden, A., Yanuka, O., Amit, M., Soreq, H., and Benvenisty, N. (2000a). Differentiation of human embryonic stem cells into embryoid bodies comprising the three embryonic germ layers. Mol Med *6*, 88-95.

Itskovitz-Eldor, J., Schuldiner, M., Karsenti, D., Eden, A., Yanuka, O., Amit, M., Soreq, H., and Benvenisty, N. (2000b). Differentiation of human embryonic stem cells into embryoid bodies compromising the three embryonic germ layers. Mol Med *6*, 88-95.

Jego, G., Hazoume, A., Seigneuric, R., and Garrido, C. (2013). Targeting heat shock proteins in cancer. Cancer Lett *332*, 275-285.

Josephson, R., Ording, C.J., Liu, Y., Shin, S., Lakshmipathy, U., Toumadje, A., Love, B., Chesnut, J.D., Andrews, P.W., Rao, M.S., *et al.* (2007). Qualification of embryonal carcinoma 2102Ep as a reference for human embryonic stem cell research. Stem Cells *25*, 437-446.

Kane, N.M., Meloni, M., Spencer, H.L., Craig, M.A., Strehl, R., Milligan, G., Houslay, M.D., Mountford, J.C., Emanueli, C., and Baker, A.H. (2010). Derivation of endothelial cells from human embryonic stem cells by directed differentiation: analysis of microRNA and angiogenesis in vitro and in vivo. Arterioscler Thromb Vasc Biol *30*, 1389-1397.

Kennedy, M., D'Souza, S.L., Lynch-Kattman, M., Schwantz, S., and Keller, G. (2007). Development of the hemangioblast defines the onset of hematopoiesis in human ES cell differentiation cultures. Blood *109*, 2679-2687.

Kim, D.S., Lee, J.S., Leem, J.W., Huh, Y.J., Kim, J.Y., Kim, H.S., Park, I.H., Daley, G.Q., Hwang, D.Y., and Kim, D.W. (2010). Robust Enhancement of Neural

Differentiation from Human ES and iPS Cells Regardless of their Innate Difference in Differentiation Propensity. Stem Cell Rev Rep *6*, 270-281.

Kono, T., Kubo, H., Shimazu, C., Ueda, Y., Takahashi, M., Yanagi, K., Fujita, N., Tsuruo, T., Wada, H., and Yamashita, J.K. (2006). Differentiation of lymphatic endothelial cells from embryonic stem cells on OP9 stromal cells. Arterioscl Throm Vas *26*, 2070-2076.

Kumagai, H., Suemori, H., Uesugi, M., Nakatsuji, N., and Kawase, E. (2013). Identification of small molecules that promote human embryonic stem cell self-renewal. Biochem Bioph Res Co *434*, 710-716.

Lancaster, M.A., Renner, M., Martin, C.A., Wenzel, D., Bicknell, L.S., Hurles, M.E., Homfray, T., Penninger, J.M., Jackson, A.P., and Knoblich, J.A. (2013). Cerebral organoids model human brain development and microcephaly. Nature *501*, 373-+.

Lee, H., Park, J., Forget, B.G., and Gaines, P. (2009). Induced pluripotent stem cells in regenerative medicine: an argument for continued research on human embryonic stem cells. Regenerative Medicine *4*, 759-769.

Leslie, B.J., and Hergenrother, P.J. (2008). Identification of the cellular targets of bioactive small organic molecules using affinity reagents. Chem Soc Rev *37*, 1347-1360.

Levenberg, S., Golub, J.S., Amit, M., Itskovitz-Eldor, J., and Langer, R. (2002). Endothelial cells derived from human embryonic stem cells. Proc Natl Acad Sci U S A *99*, 4391-4396.

Li, J., Wang, G.W., Wang, C.Y., Zhao, Y., Zhang, H., Tan, Z.J., Song, Z.H., Ding, M.X., and Deng, H.K. (2007). MEK/ERK signaling contributes to the maintenance of human embryonic stem cell self-renewal. Differentiation *75*, 299-307.

Li, W., Tian, E., Chen, Z.X., Sun, G., Ye, P., Yang, S., Lu, D., Xie, J., Ho, T.V., Tsark, W.M.*, et al.* (2012). Identification of Oct4-activating compounds that enhance reprogramming efficiency. P Natl Acad Sci USA *109*, 20853-20858.

Li, W.L., Zhou, H.Y., Abujarour, R., Zhu, S.Y., Joo, J.Y., Lin, T.X., Hao, E.G., Scholer, H.R., Hayek, A., and Ding, S. (2009). Generation of Human-Induced Pluripotent Stem Cells in the Absence of Exogenous Sox2. Stem Cells *27*, 2992-3000.

Lian, X.J., Hsiao, C., Wilson, G., Zhu, K.X., Hazeltine, L.B., Azarin, S.M., Raval, K.K., Zhang, J.H., Kamp, T.J., and Palecek, S.P. (2012). Robust cardiomyocyte differentiation from human pluripotent stem cells via temporal modulation of canonical Wnt signaling. P Natl Acad Sci USA *109*, E1848-E1857.

Lin, T., Ambasudhan, R., Yuan, X., Li, W., Hilcove, S., Abujarour, R., Lin, X., Hahm, H.S., Hao, E., Hayek, A.*, et al.* (2009). A chemical platform for improved induction of human iPSCs. Nat Methods *6*, 805-808.

Lomenick, B., Hao, R., Jonai, N., Chin, R.M., Aghajan, M., Warburton, S., Wang, J.N., Wu, R.P., Gomez, F., Loo, J.A., *et al.* (2009). Target identification using drug affinity responsive target stability (DARTS). P Natl Acad Sci USA *106*, 21984-21989.

Lomenick, B., Olsen, R.W., and Huang, J. (2011). Identification of Direct Protein Targets of Small Molecules. Acs Chem Biol *6*, 34-46.

Ludwig, T.E., Levenstein, M.E., Jones, J.M., Berggren, W.T., Mitchen, E.R., Frane, J.L., Crandall, L.J., Daigh, C.A., Conard, K.R., Piekarczyk, M.S., *et al.* (2006). Derivation of human embryonic stem cells in defined conditions. Nat Biotechnol *24*, 185-187.

Macarron, R., Banks, M.N., Bojanic, D., Burns, D.J., Cirovic, D.A., Garyantes, T., Green, D.V.S., Hertzberg, R.P., Janzen, W.P., Paslay, J.W., *et al.* (2011). Impact of high-throughput screening in biomedical research. Nat Rev Drug Discov *10*, 188-195.

Macielag, M.J. (2012). Chemical Properties of Antimicrobials and Their Uniqueness. Antibiotic Discovery and Development, Vols 1 and 2, 793-820.

Maherali, N., and Hochedlinger, K. (2009). Tgf beta Signal Inhibition Cooperates in the Induction of iPSCs and Replaces Sox2 and cMyc. Curr Biol *19*, 1718-1723.

Mahmood, A., Harkness, L., Schroder, H.D., Abdallah, B.M., and Kassem, M. (2010). Enhanced Differentiation of Human Embryonic Stem Cells to Mesenchymal Progenitors by Inhibition of TGF-beta/Activin/Nodal Signaling Using SB-431542. J Bone Miner Res *25*, 1216-1233.

Mak, S.K., Huang, Y.A., Iranmanesh, S., Vangipuram, M., Sundararajan, R., Nguyen, L., Langston, J.W., and Schule, B. (2012). Small Molecules Greatly Improve Conversion of Human-Induced Pluripotent Stem Cells to the Neuronal Lineage. Stem Cells Int.

Mali, P., Chou, B.K., Yen, J., Ye, Z., Zou, J., Dowey, S., Brodsky, R.A., Ohm, J.E., Yu, W., Baylin, S.B., *et al.* (2010). Butyrate greatly enhances derivation of human induced pluripotent stem cells by promoting epigenetic remodeling and the expression of pluripotency-associated genes. Stem Cells *28*, 713-720.

Mariani, J., Simonini, M.V., Palejev, D., Tomasini, L., Coppola, G., Szekely, A.M., Horvath, T.L., and Vaccarino, F.M. (2012). Modeling human cortical development in vitro using induced pluripotent stem cells. P Natl Acad Sci USA *109*, 12770-12775.

Martinez-Corral, I., and Makinen, T. (2013). Regulation of lymphatic vascular morphogenesis: Implications for pathological (tumor) lymphangiogenesis. Experimental Cell Research *319*, 1618-1625.

Massey, A.J., Williamson, D.S., Browne, H., Murray, J.B., Dokurno, P., Shaw, T., Macias, A.T., Daniels, Z., Geoffroy, S., Dopson, M., *et al.* (2010). A novel, small molecule inhibitor of Hsc70/Hsp70 potentiates Hsp90 inhibitor induced apoptosis in HCT116 colon carcinoma cells. Cancer Chemother Pharmacol *66*, 535-545.

McCracken, K.W., Cata, E.M., Crawford, C.M., Sinagoga, K.L., Schumacher, M., Rockich, B.E., Tsai, Y.H., Mayhew, C.N., Spence, J.R., Zavros, Y., *et al.* (2014). Modelling human development and disease in pluripotent stem-cell-derived gastric organoids. Nature *516*, 400-404.

McLean, A.B., D'Amour, K.A., Jones, K.L., Krishnamoorthy, M., Kulik, M.J., Reynolds, D.M., Sheppard, A.M., Liu, H.Q., Xu, Y., Baetge, E.E., *et al.* (2007). Activin a efficiently specifies definitive endoderm from human embryonic stem cells only when phosphatidylinositol 3-kinase signaling is suppressed. Stem Cells *25*, 29-38.

Moschidou, D., Mukherjee, S., Blundell, M.P., Drews, K., Jones, G.N., Abdulrazzak, H., Nowakowska, B., Phoolchund, A., Lay, K., Ramasamy, T.S., *et al.* (2012). Valproic Acid Confers Functional Pluripotency to Human Amniotic Fluid Stem Cells in a Transgene-free Approach. Mol Ther *20*, 1953-1967.

Muguruma, K., Nishiyama, A., Ono, Y., Miyawaki, H., Mizuhara, E., Hori, S., Kakizuka, A., Obata, K., Yanagawa, Y., Hirano, T., *et al.* (2010). Ontogeny-recapitulating generation and tissue integration of ES cell-derived Purkinje cells. Nat Neurosci *13*, 1171-1180.

Muller, L., Schaupp, A., Walerych, D., Wegele, H., and Buchner, J. (2004). Hsp90 regulates the activity of wild type p53 under physiological and elevated temperatures. Journal of Biological Chemistry *279*, 48846-48854.

Nair, A.V., Hocher, B., Verkaart, S., van Zeeland, F., Pfab, T., Slowinski, T., Chen, Y.P., Schlingmann, K.P., Schaller, A., Gallati, S., *et al.* (2012). Loss of insulin-induced activation of TRPM6 magnesium channels results in impaired glucose tolerance during pregnancy. P Natl Acad Sci USA *109*, 11324-11329.

Nakajima-Takagi, Y., Osawa, M., Oshima, M., Takagi, H., Miyagi, S., Endoh, M., Endo, T.A., Takayama, N., Eto, K., Toyoda, T., *et al.* (2013). Role of SOX17 in hematopoietic development from human embryonic stem cells. Blood *121*, 447-458.

Nakano, T., Ando, S., Takata, N., Kawada, M., Muguruma, K., Sekiguchi, K., Saito, K., Yonemura, S., Eiraku, M., and Sasai, Y. (2012). Self-formation of optic cups and storable stratified neural retina from human ESCs. Cell Stem Cell *10*, 771-785.

Ng, H.H., and Surani, M.A. (2011). The transcriptional and signalling networks of pluripotency. Nat Cell Biol *13*, 490-496.

Norrman, K., Strombeck, A., Semb, H., and Stahlberg, A. (2013). Distinct gene expression signatures in human embryonic stem cells differentiated towards definitive endoderm at single-cell level. Methods *59*, 59-70.

Passier, R. (2003). Potential of human embryonic stem cells in regenerative medicine. Horm Res *60*, 11-14.

Perrier, A.L., Tabar, V., Barberi, T., Rubio, M.E., Bruses, J., Topf, N., Harrison, N.L., and Studer, L. (2004). Derivation of midbrain dopamine neurons from human embryonic stem cells. Proc Natl Acad Sci U S A *101*, 12543-12548.

Pohl, M., Stuart, R.O., Sakurai, H., and Nigam, S.K. (2000). Branching morphogenesis during kidney development. Annu Rev Physiol *62*, 595-620.

Rodda, D.J., Chew, J.L., Lim, L.H., Loh, Y.H., Wang, B., Ng, H.H., and Robson, P. (2005). Transcriptional regulation of nanog by OCT4 and SOX2. J Biol Chem *280*, 24731-24737.

Rosado, A., Sohn, E.J., Drakakaki, G., Pan, S.Q., Swidergal, A., Xiong, Y.Q., Kang, B.H., Bressan, R.A., and Raikhel, N.V. (2010). Auxin-Mediated Ribosomal Biogenesis Regulates Vacuolar Trafficking in Arabidopsis. Plant Cell *22*, 143-158.

Ryazanova, L.V., Rondon, L.J., Zierler, S., Hu, Z.X., Galli, J., Yamaguchi, T.P., Mazur, A., Fleig, A., and Ryazanov, A.G. (2010). TRPM7 is essential for Mg2+ homeostasis in mammals. Nat Commun *1*.

Saito, H., Takeuchi, M., Chida, K., and Miyajima, A. (2011). Generation of glucose-responsive functional islets with a three-dimensional structure from mouse fetal pancreatic cells and iPS cells in vitro. Plos One *6*, e28209.

Salven, P., Mustjoki, S., Alitalo, R., Alitalo, K., and Rafii, S. (2003). VEGFR-3 and CD133 identify a population of CD34(+) lymphatic/vascular endothelial precursor cells. Blood *101*, 168-172.

Samuel, R., Daheron, L., Liao, S., Vardam, T., Kamoun, W.S., Batista, A., Buecker, C., Schafer, R., Han, X., Au, P.*, et al.* (2013). Generation of functionally competent and durable engineered blood vessels from human induced pluripotent stem cells. Proc Natl Acad Sci U S A *110*, 12774-12779.

Sasai, Y. (2013). Next-generation regenerative medicine: organogenesis from stem cells in 3D culture. Cell Stem Cell *12*, 520-530.

Schenone, M., Dancik, V., Wagner, B.K., and Clemons, P.A. (2013). Target identification and mechanism of action in chemical biology and drug discovery. Nature Chemical Biology *9*, 232-240.

Schlingmann, K.P., Waldegger, S., Konrad, M., Chubanov, V., and Gudermann, T. (2007). TRPM6 and TRPM7 - Gatekeepers of human magnesium metabolism. Bba-Mol Basis Dis *1772*, 813-821.

Schulte-Merker, S., Sabine, A., and Petrova, T.V. (2011). Lymphatic vascular morphogenesis in development, physiology, and disease. Journal of Cell Biology *193*, 607-618.

Shaknovich, R., Shue, G., and Kohtz, D.S. (1992). Conformational activation of a basic helix-loop-helix protein (MyoD1) by the C-terminal region of murine HSP90 (HSP84). Mol Cell Biol *12*, 5059-5068.

Shu, J., Wu, C., Wu, Y.T., Li, Z.Y., Shao, S.D., Zhao, W.H., Tang, X., Yang, H., Shen, L.J., Zuo, X.H., *et al.* (2013). Induction of Pluripotency in Mouse Somatic Cells with Lineage Specifiers. Cell *153*, 963-975.

Son, Y.S., Park, J.H., Kang, Y.K., Park, J.S., Choi, H.S., Lim, J.Y., Lee, J.E., Lee, J.B., Ko, M.S., Kim, Y.S., *et al.* (2005). Heat shock 70-kDa protein 8 isoform 1 is expressed on the surface of human embryonic stem cells and downregulated upon differentiation. Stem Cells *23*, 1502-1513.

Sone, M., Itoh, H., Yamahara, K., Yamashita, J.K., Yurugi-Kobayashi, T., Nonoguchi, A., Suzuki, Y., Chao, T.H., Sawada, N., Fukunaga, Y., *et al.* (2007). Pathway for differentiation of human embryonic stem cells to vascular cell components and their potential for vascular regeneration. Arterioscl Throm Vas *27*, 2127-2134.

Spence, J.R., Mayhew, C.N., Rankin, S.A., Kuhar, M.F., Vallance, J.E., Tolle, K., Hoskins, E.E., Kalinichenko, V.V., Wells, S.I., Zorn, A.M., *et al.* (2011). Directed differentiation of human pluripotent stem cells into intestinal tissue in vitro. Nature *470*, 105-U120.

Stavreva, D.A., Muller, W.G., Hager, G.L., Smith, C.L., and McNally, J.G. (2004). Rapid glucocorticoid receptor exchange at a promoter is coupled to transcription and regulated by chaperones and proteasomes. Molecular and Cellular Biology *24*, 2682-2697.

Sternlicht, M.D. (2006). Key stages in mammary gland development - The cues that regulate ductal branching morphogenesis. Breast Cancer Res *8*.

Stricher, F., Macri, C., Ruff, M., and Muller, S. (2013). HSPA8/HSC70 chaperone protein: structure, function, and chemical targeting. Autophagy *9*, 1937-1954.

Suga, H., Kadoshima, T., Minaguchi, M., Ohgushi, M., Soen, M., Nakano, T., Takata, N., Wataya, T., Muguruma, K., Miyoshi, H., *et al.* (2011). Self-formation of functional adenohypophysis in three-dimensional culture. Nature *480*, 57-62.

Suzuki, H., Watabe, T., Kato, M., Miyazawa, K., and Miyazono, K. (2005). Roles of vascular endothelial growth factor receptor 3 signaling in differentiation of mouse embryonic stem cell-derived vascular progenitor cells into endothelial cells. Blood *105*, 2372-2379.

Swaminathan, R. (2003). Magnesium metabolism and its disorders. Clin Biochem Rev *24*, 47-66.

Tahamtani, Y., Azarnia, M., Farrokhi, A., Sharifi-Zarchi, A., Aghdami, N., and Baharvand, H. (2013). Treatment of Human Embryonic Stem Cells with Different

Combinations of Priming and Inducing Factors Toward Definitive Endoderm. Stem Cells Dev *22*, 1419-1432.

Takasato, M., Er, P.X., Becroft, M., Vanslambrouck, J.M., Stanley, E.G., Elefanty, A.G., and Little, M.H. (2014). Directing human embryonic stem cell differentiation towards a renal lineage generates a self-organizing kidney. Nat Cell Biol *16*, 118-126.

Tammela, T., and Alitalo, K. (2010). Lymphangiogenesis: Molecular Mechanisms and Future Promise. Cell *140*, 460-476.

Tan, Y.Z., Wang, H.J., Zhang, M.H., Quan, Z., Li, T., and He, Q.Z. (2014). CD34(+)VEGFR-3(+) progenitor cells have a potential to differentiate towards lymphatic endothelial cells. J Cell Mol Med *18*, 422-433.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. In Science, pp. 1145-1147.

Tsutsui, H., Valamehr, B., Hindoyan, A., Qiao, R., Ding, X.T., Guo, S.L., Witte, O.N., Liu, X., Ho, C.M., and Wu, H. (2011). An optimized small molecule inhibitor cocktail supports long-term maintenance of human embryonic stem cells. Nat Commun *2*.

Ueda, T., Yamada, T., Hokuto, D., Koyama, F., Kasuda, S., Kanehiro, H., and Nakajima, Y. (2010). Generation of functional gut-like organ from mouse induced pluripotent stem cells. Biochem Biophys Res Commun *391*, 38-42.

Vallier, L., Alexander, M., and Pedersen, R.A. (2005). Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. J Cell Sci *118*, 4495-4509.

Vestweber, D. (2008). VE-cadherin - The major endothelial adhesion molecule controlling cellular junctions and blood vessel formation. Arterioscl Throm Vas *28*, 223-232.

Vodyanik, M.A., Thomson, J.A., and Slukvin, I.I. (2006). Leukosialin (CD43) defines hematopoietic progenitors in human embryonic stem cell differentiation cultures. Blood *108*, 2095-2105.

Voyta, J.C., Via, D.P., Butterfield, C.E., and Zetter, B.R. (1984). Identification and Isolation of Endothelial-Cells Based on Their Increased Uptake of Acetylated-Low Density Lipoprotein. J Cell Biol *99*, 2034-2040.

Walder, R.Y., Yang, B.L., Stokes, J.B., Kirby, P.A., Cao, X., Shi, P.J., Searby, C.C., Husted, R.F., and Sheffield, V.C. (2009). Mice defective in Trpm6 show embryonic mortality and neural tube defects. Hum Mol Genet *18*, 4367-4375.

Walerych, D., Kudla, G., Gutkowska, M., Wawrzynow, B., Muller, L., King, F.W., Helwak, A., Boros, J., Zylicz, A., and Zylicz, M. (2004). Hsp90 chaperones wild-type p53 tumor suppressor protein. Journal of Biological Chemistry *279*, 48836-48845.

Wang, Q., Xu, X.X., Li, J., Liu, J., Gu, H.F., Zhang, R., Chen, J.K., Kuang, Y., Fei, J., Jiang, C., *et al.* (2011). Lithium, an anti-psychotic drug, greatly enhances the generation of induced pluripotent stem cells. Cell Res *21*, 1424-1435.

Wang, X.L., and Ye, K.M. (2009). Three-Dimensional Differentiation of Embryonic Stem Cells into Islet-like Insulin-Producing Clusters. Tissue Eng Pt A *15*, 1941-1952.

Wang, Z.Z., Au, P., Chen, T., Shao, Y., Daheron, L.M., Bai, H., Arzigian, M., Fukumura, D., Jain, R.K., and Scadden, D.T. (2007). Endothelial cells derived from human embryonic stem cells form durable blood vessels in vivo. Nat Biotechnol *25*, 317-318.

Ware, C.B., Wang, L.L., Mecham, B.H., Shen, L.L., Nelson, A.M., Bar, M., Lamba, D.A., Dauphin, D.S., Buckingham, B., Askari, B., *et al.* (2009). Histone Deacetylase Inhibition Elicits an Evolutionarily Conserved Self-Renewal Program in Embryonic Stem Cells. Cell Stem Cell *4*, 359-369.

Watanabe, K., Ueno, M., Kamiya, D., Nishiyama, A., Matsumura, M., Wataya, T., Takahashi, J.B., Nishikawa, S., Nishikawa, S., Muguruma, K., *et al.* (2007). A ROCK inhibitor permits survival of dissociated human embryonic stem cells. Nat Biotechnol *25*, 681-686.

Wegele, H., Muller, L., and Buchner, J. (2004). Hsp70 and Hsp90 - a relay team for protein folding. Rev Physiol Bioch P *151*, 1-44.

Wei, X., Chen, Y., Xu, Y., Zhan, Y., Zhang, R., Wang, M., Hua, Q., Gu, H., Nan, F., and Xie, X. (2014). Small molecule compound induces chromatin de-condensation and facilitates induced pluripotent stem cell generation. J Mol Cell Biol *6*, 409-420.

Wells, J.M., and Spence, J.R. (2014). How to make an intestine. Development *141*, 752-760.

Wolf, F.I., and Trapani, V. (2012). Magnesium and its transporters in cancer: a novel paradigm in tumour development. Clin Sci *123*, 417-427.

Wong, A.P., Bear, C.E., Chin, S., Pasceri, P., Thompson, T.O., Huan, L.J., Ratjen, F., Ellis, J., and Rossant, J. (2012). Directed differentiation of human pluripotent stem cells into mature airway epithelia expressing functional CFTR protein. Nature Biotechnology *30*, 876-U108.

Wu, S.M., and Hothedlinger, K. (2011). Harnessing the potential of induced pluripotent stem cells for regenerative medicine. Nat Cell Biol *13*, 497-505.

Xia, Y., Nivet, E., Sancho-Martinez, I., Gallegos, T., Suzuki, K., Okamura, D., Wu, M.Z., Dubova, I., Esteban, C.R., Montserrat, N., *et al.* (2013). Directed differentiation of human pluripotent cells to ureteric bud kidney progenitor-like cells. Nat Cell Biol *15*, 1507-+.

Xu, C., Inokuma, M.S., Denham, J., Golds, K., Kundu, P., Gold, J.D., and Carpenter, M.K. (2001). Feeder-free growth of undifferentiated human embryonic stem cells. Nat Biotechnol *19*, 971-974.

Xu, R.H., Chen, X., Li, D.S., Li, R., Addicks, G.C., Glennon, C., Zwaka, T.P., and Thomson, J.A. (2002). BMP4 initiates human embryonic stem cell differentiation to trophoblast. Nature Biotechnology *20*, 1261-1264.

Xu, Y., Shi, Y., and Ding, S. (2008). A chemical approach to stem-cell biology and regenerative medicine. Nature *453*, 338-344.

Xu, Y., Zhu, X., Hahm, H.S., Wei, W., Hao, E., Hayek, A., and Ding, S. (2010a). Revealing a core signaling regulatory mechanism for pluripotent stem cell survival and self-renewal by small molecules. Proc Natl Acad Sci U S A *107*, 8129-8134.

Xu, Y., Zhu, X.W., Hahm, H.S., Wei, W.G., Hao, E.G., Hayek, A., and Ding, S. (2010b). Revealing a core signaling regulatory mechanism for pluripotent stem cell survival and self-renewal by small molecules. P Natl Acad Sci USA *107*, 8129-8134.

Yabut, O., and Bernstein, H.S. (2011). Human Embryonic Stem Cells in Regenerative Medicine. Stem Cells Biol Reg, 17-38.

Yen, J., Yin, L., and Cheng, J. (2014). Enhanced Non-Viral Gene Delivery to Human Embryonic Stem Cells via Small Molecule-Mediated Transient Alteration of Cell Structure. J Mater Chem B Mater Biol Med *2*, 8098-8105.

Ying, Q.L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. Nature *453*, 519-U515.

Yu, C., Liu, K., Tang, S., and Ding, S. (2014). Chemical approaches to cell reprogramming. Curr Opin Genet Dev *28C*, 50-56.

Yu, P., Pan, G., Yu, J., and Thomson, J.A. (2011). FGF2 sustains NANOG and switches the outcome of BMP4-induced human embryonic stem cell differentiation. Cell Stem Cell *8*, 326-334.

Yuan, X., Wan, H.F., Zhao, X.Y., Zhu, S.Y., Zhou, Q., and Ding, S. (2011). Brief Report: Combined Chemical Treatment Enables Oct4-Induced Reprogramming from Mouse Embryonic Fibroblasts. Stem Cells *29*, 549-553.

Zhang, S.C., Wernig, M., Duncan, I.D., Brustle, O., and Thomson, J.A. (2001). In vitro differentiation of transplantable neural precursors from human embryonic stem cells. Nat Biotechnol *19*, 1129-1133.

Zhang, Z., Yu, H.J., Huang, J.H., Faouzi, M., Schmitz, C., Penner, R., and Fleig, A. (2014). The TRPM6 Kinase Domain Determines the Mg center dot ATP Sensitivity of TRPM7/M6 Heteromeric Ion Channels*. Journal of Biological Chemistry *289*, 5217-5227.

Zhou, H.Y., Li, W.L., Zhu, S.Y., Joo, J.Y., Do, J.T., Xiong, W., Kim, J.B., Zhang, K., Scholer, H.R., and Ding, S. (2010a). Conversion of Mouse Epiblast Stem Cells to an Earlier Pluripotency State by Small Molecules. Journal of Biological Chemistry *285*, 29676-29680.

Zhou, J., Su, P., Li, D., Tsang, S., Duan, E., and Wang, F. (2010b). High-efficiency induction of neural conversion in human ESCs and human induced pluripotent stem cells with a single chemical inhibitor of transforming growth factor beta superfamily receptors. Stem Cells *28*, 1741-1750.

Zhu, S., Wurdak, H., Wang, J., Lyssiotis, C.A., Peters, E.C., Cho, C.Y., Wu, X., and Schultz, P.G. (2009). A small molecule primes embryonic stem cells for differentiation. Cell Stem Cell *4*, 416-426.

Zhu, S.Y., Li, W.L., Zhou, H.Y., Wei, W.G., Ambasudhan, R., Lin, T.X., Kim, J., Zhang, K., and Ding, S. (2010). Reprogramming of Human Primary Somatic Cells by OCT4 and Chemical Compounds. Cell Stem Cell *7*, 651-655.

Ziegler, S., Pries, V., Hedberg, C., and Waldmann, H. (2013). Target Identification for Small Bioactive Molecules: Finding the Needle in the Haystack. Angew Chem Int Edit *52*, 2744-2792.

zur Nieden, N.I., Kempka, G., Rancourt, D.E., and Ahr, H.J. (2005). Induction of chondro-, osteo- and adipogenesis in embryonic stem cells by bone morphogenetic protein-2: Effect of cofactors on differentiating lineages. Bmc Dev Biol *5*.